

161444

Shears, Beverly

From: Devi, Sarvamangala
Sent: Wednesday, July 27, 2005 3:59 PM
To: Shears, Beverly
Subject: 10/070,882

Beverly:

In application 10/070,882, please perform a sequence search for SEQ ID NO: 2 in commercial and pending databases.

Thanx.

S. DEVI, Ph.D.
Primary Examiner
AU 1645
Rems - 3C18

UA
Pending
Next

1

Date completed:
Searcher: Beverly e 2528
Terminal time:
Elapsed time:
CPU time:
Total time:
Number of Searches:
Number of Databases:

Search Site	Vendors
<input type="checkbox"/> STIC	<input type="checkbox"/> IG
<input type="checkbox"/> CM-1	<input type="checkbox"/> STN
<input type="checkbox"/> Pre-S	<input type="checkbox"/> Dialog
Type of Search	
<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> APS
<input type="checkbox"/> A.A. Sequence	<input type="checkbox"/> Geninfo
<input type="checkbox"/> Structure	<input type="checkbox"/> SDC
<input type="checkbox"/> Bibliographic	<input checked="" type="checkbox"/> DARC/Questel
	<input type="checkbox"/> Other CGN

This Page Blank (uspto)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 1051-144

TO: Sarvamangala Devi
Art Unit: 1645
Location: REM 3C18
Serial Number: 10/070882

Tuesday, August 09, 2005

From: Beverly Shears
Location: Biotech-Chem Library
REM 1A54
Phone: 571-272-2528
beverly.shears@uspto.gov

Search Notes

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Q_b nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:30:55 ; Search time 7070 Seconds

US-10-070-882a-2
Perfect score: 139
Sequence: 1 gtagacttggcgacgaaact.....ctgttcttattgttaacaca 139

748.364 Million cell updates/sec

Title: US-10-070-882a-2

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3423954 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST*

1: gb_efst1:*

2: gb_efst2:*

3: gb_hrc:*

4: gb_efst3:*

5: gb_efst4:*

6: gb_efst5:*

7: gb_efst6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
c 1	34.8	25.0	3332	AK082454
c 2	34.4	24.7	414	B1642267
c 3	33.6	24.2	204	AV231406
c 4	33.6	24.2	274	B8775700
c 5	33.6	24.2	318	B8094276
c 6	33.6	24.2	335	BB222962
c 7	33.6	24.2	448	BY439440
c 8	33.6	24.2	570	A2435636
c 9	33.6	24.2	590	BR131381
c 10	33.6	24.2	605	BM0598097
c 11	33.6	24.2	625	CB835622
c 12	33.6	24.2	633	CA834105
c 13	33.6	24.2	674	BM01497
c 14	33.6	24.2	699	BM301422
c 15	33.6	24.2	746	CR303205
c 16	33.6	24.2	863	BR035174
c 17	33.6	24.2	866	CR308064
c 18	33.6	24.2	889	CG92370
c 19	33.6	24.2	1192	BE03711
c 20	33.4	24.0	521	AO830244
c 21	33.4	24.0	728	MUS_muscul
c 22	33.4	24.0	957	CA791309
c 23	33.2	23.9	205	BR004186
c 24	33.2	23.9	467	BZ915818

ALIGNMENTS
RESULT 1 AK082454/c
LOCUS AK082454
DEFINITION Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched library, clone:C230052F14 product:unclassifiable, full insert sequence.
ACCESSION AK082454
VERSION AK082454.1 GI:26100683
KEYWORDS HTC; CAP-trapper.
SOURCE
ORGANISM Mus musculus (house mouse)
ORGANISM Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 Carninci, P. and Hayashizaki, Y. Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253 10349656
REFERENCE
AUTHORS Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20439374
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Tachiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, I., Ikegami, T., Kasinagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataiaki, M., Yonetani, Y., Ichikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inouye, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20550913
PUBMED 11076861
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
FUNCTIONAL annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research

	SOURCE
JOURNAL	ORGANISM
TITLE	Group Phase I & II team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 3932)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furukubo,M., Hanagaki,T., Hara,A., Haehizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirokane,T., Hiraoka,T., Hirono,I., Kasukawa,T., Kato,J., Kawai,Y., Kojima,Y., Kondo,S., Konno,H., Kozai,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasada,D., Shirota,K., Shinagawa,A., Shiraki,T., Socabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Tomaru,A., Toyoda,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration, Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-reseq@riken.jp, URL: http://fantom.gsc.riken.jp/)
COMMENT	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration, Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-reseq@riken.jp, URL: http://fantom.gsc.riken.jp/)
FEATURES	Location/Qualifiers
source	1. J.3932 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:C230052F14" /db_xref="taxon:10990" /clones="C230052F14" /tissue="cerebellum" /clone_Tiba="RIKEN full-length enriched mouse cDNA library" /dev_BTstage="0 day neonate" /note="unclassifiable" 1. J.3932 /note="unclassifiable"
ORIGIN	Query Match 25.0% Score 34.8; DB 3; length 3932; Best Local Similarity 54.8%; Pred. No. 13; Mismatches 57; Indels 0; Gaps 0; Matches 69; Conservative
QY	14 AGCAACTAATAATGCCTGCTCACCTCTTCTCAGAGAGGGTGACTATTGTC 73 Db 2194 ACTATCTCATTAAGCTCTCTATGTGTGTTAATGTTGGATTTG 2135
QY	74 TCGTTTAACTACTGTATCCCAAAGCACCATAATACCGTAGCTGTCTTATGTT 133 Db 2134 TGGTATATGACTGATTCTCAGATCACATTGATAAATGATGATGTCAGTGT 2075
QY	134 AACACA 139 Db 2074 CAGCA 2069
RESULT 2	BY642267 LOCUS BY642267 RIKEN full-length enriched, visual cortex Mus musculus DEFINITION cDNA Clone K430350A22 3', mRNA sequence. ACCESSION BY642267 VERSION BY642267.1 KEYWORDS EST.
COMMENT	414 bp mRNA linear EST 15-DEC-2002 Tissues were provided by Michela Fagiolini and Takeo K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiyosawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
JOURNAL	Mus musculus (house mouse)
REFERENCE	Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamamoto,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gotohori,T., Baladreli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbin,L.E., Cousins,S., Dalia,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Fraser,K.S., Gaasterland,T., Garibaldi,M., Glasb,C., Godzik,A., Gough,J., Grimmond,S., Guatamacich,S., Hirokawa,N., Jacobson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawabaya,Y., Kadzienski,R.M., King,B.L., Kongyaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maitais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavani,W.J., Pertea,G., Peacock,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramchandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watabayashi,Y., Wells,C., Wilming,L.G., Wymshaw,Boris,A., Yangisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirokane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Wakai,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kadowa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibusawa,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,B. and Hayashizaki,Y.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
COMMENT	Contact: Yoshihide Hayashizaki, Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute, The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-803-9222 Fax: 81-45-503-9216 Email: genome-reseq@riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirokane,T., Imotani,K., Ishii,Y., Ichoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Watanuki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedias: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Encyclopedias Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES	Location/Qualifiers
source	1. J.3932 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:C230052F14" /db_xref="taxon:10990" /clones="C230052F14" /tissue="cerebellum" /clone_Tiba="RIKEN full-length enriched mouse cDNA library" /dev_BTstage="0 day neonate" /note="unclassifiable" 1. J.3932 /note="unclassifiable"
ORIGIN	Query Match 25.0% Score 34.8; DB 3; length 3932; Best Local Similarity 54.8%; Pred. No. 13; Mismatches 57; Indels 0; Gaps 0; Matches 69; Conservative
QY	14 AGCAACTAATAATGCCTGCTCACCTCTTCTCAGAGAGGGTGACTATTGTC 73 Db 2194 ACTATCTCATTAAGCTCTCTATGTGTGTTAATGTTGGATTTG 2135
QY	74 TCGTTTAACTACTGTATCCCAAAGCACCATAATACCGTAGCTGTCTTATGTT 133 Db 2134 TGGTATATGACTGATTCTCAGATCACATTGATAAATGATGATGTCAGTGT 2075
Db	134 AACACA 139 2074 CAGCA 2069
RESULT 2	BY642267 LOCUS BY642267 RIKEN full-length enriched, visual cortex Mus musculus DEFINITION cDNA Clone K430350A22 3', mRNA sequence. ACCESSION BY642267 VERSION BY642267.1 KEYWORDS EST.
COMMENT	414 bp mRNA linear EST 15-DEC-2002 Tissues were provided by Michela Fagiolini and Takeo K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiyosawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

	<p>Tel: 81-45-503-9222 Email: genome-res@sc.riken.jp, url: http://genome.gsc.riken.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. <i>Genome Res.</i> . 10 (10), 1611-1630 (2000) "wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Kono, H., Fukunishi, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Itoh, M., Carninci, P., Muramatsu, M., Kawai, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. Sequence analysis (RSEA) system -384 format sequencing pipeline with 384 multicapillary sequencer. <i>Genome Res.</i> . 10 (11), 1757-1771 (2000)</p> <p>Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. <i>Genome Res.</i> . 11 (2), 281-289 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details.</p>												
FEATURES	<p>Location/Qualifiers</p> <p>1. 274 <i>/organism="Mus musculus"</i> <i>/mol_type="mRNA"</i> <i>/strain="DD2"</i> <i>/db_xref="taxon:10090"</i> <i>/clone="G430026109"</i> <i>/cell_line="RCB-0559 K-1.F1"</i> <i>/clone_lip="RIKEN full-length enriched, RCB-0559 K-1 . F1</i> <i>/note="Organ: spleen"</i></p>												
FEATURES	<p>source</p> <p>1. 318 <i>/organism="Mus musculus"</i> <i>/mol_type="mRNA"</i> <i>/db_xref="taxon:10090"</i> <i>/clone="G430045021"</i> <i>/tissue_type="embryonic body between diaphragm region and neck"</i> <i>/dev_stage="12 days embryo"</i> <i>/lab_host="DH10B"</i> <i>/clone_lip="RIKEN full-length enriched, 12 days embryo, embryonic body between diaphragm region and neck"</i> <i>/note="Site 1; Sali; Site 2; BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'- GAGAGAGAGAGATCCAGAGCTTCTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the Primer adapter of sequence [5' GAGAGAGAGATCTCGAGTAATTAACTCCCCCCCC 3']. cDNA was cleaved with XbaI and BamHI."</i></p>												
ORIGIN	<p>Query Match</p> <p>Best Local Similarity 57.7%; Pred. No. 20; Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;</p> <table border="0"> <tr> <td>Qy</td> <td>31 CTGCCCTCACCCCTTCTTCAGAAGAGGGGACTATTTGCTGGTTTATACTGTT</td> <td>90</td> </tr> <tr> <td>Db</td> <td>122 CTGCTCAGGAGCTCTCTTGAGCTGCTCCACAGTGTCAAGTTTCACTGTT</td> <td>181</td> </tr> <tr> <td>Qy</td> <td>91 ATCCCCAACCCATATCAAGCTAGCTGTCTATGTTA</td> <td>134</td> </tr> <tr> <td>Db</td> <td>182 TTCTCAGGACACCCACATGTGCTGACAGCCTGTTGCA</td> <td>225</td> </tr> </table>	Qy	31 CTGCCCTCACCCCTTCTTCAGAAGAGGGGACTATTTGCTGGTTTATACTGTT	90	Db	122 CTGCTCAGGAGCTCTCTTGAGCTGCTCCACAGTGTCAAGTTTCACTGTT	181	Qy	91 ATCCCCAACCCATATCAAGCTAGCTGTCTATGTTA	134	Db	182 TTCTCAGGACACCCACATGTGCTGACAGCCTGTTGCA	225
Qy	31 CTGCCCTCACCCCTTCTTCAGAAGAGGGGACTATTTGCTGGTTTATACTGTT	90											
Db	122 CTGCTCAGGAGCTCTCTTGAGCTGCTCCACAGTGTCAAGTTTCACTGTT	181											
Qy	91 ATCCCCAACCCATATCAAGCTAGCTGTCTATGTTA	134											
Db	182 TTCTCAGGACACCCACATGTGCTGACAGCCTGTTGCA	225											

Locus BB222962 335 bp mRNA linear EST 01-JUL-2000
 Definition BB222962 RIKEN full-length enriched, adult male aorta and vein-Mus
 Accession BB222962
 Version BB222962.1 GI:8891574
 Keywords EST.
 Source Mus musculus (house mouse)
 Organism Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphati; Muridae; Murinae; Mus.
 Reference 1 (bases 1 to 335)
 Authors Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishitaka,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kurokawa,H., Kojima,Y., Kondo,S., Koya,S., Kurinara,C., Kuwakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owada,T., Saito,H., Sato,C., Sato,K., Shiraki,T., Shibata,K., Shibusawa,Y., Shigemoto,Y., Shinagawa,A., Sozabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomonaga,N., Toyai,T., Tsunoda,Y., Watanuki,A., Watanabe,S., Yamamaka,T., Yamamoto,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y., unpublished (2000)
 Journal Contact: Yoshihide Hayashizaki
 Comment Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>

The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216

Reference 1 (bases 1 to 448)
 Authors Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamamoto,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Built,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Baralov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chochia,C., Corban,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Garboldi,M., Gissi,C., Godzik,A., Gong,J., Grimmond,S., Guscinich,S., Hikokawa,N., Jackson,I.J., Jarvis,B.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kediriski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Mikl,H., Nagashima,T., Numata,K., Okido,T., Pavani,W.J., Pertea,G., Pesoche,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,P.J., Reid,J.J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Seto,M., Shimada,R., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watcana,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yun,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashimoto,K., Ishii,Y., Itoh,M., Kagawa,I., Mizazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E., and Hayashizaki,Y., and vein".
 Note "Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGGAGCTTGTCTTCTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second

Origin Query Match 24.2%; Score 33.6; DB 2; Length 335;
 Definition Best Local Similarity 57.7%; Pred. No. 20; Matches 60; Conservative 0; MisMatches 44; Indels 0; Gaps 0;
 Accession BY439440
 Version BY439440.1
 Keywords EST.
 Source Mus musculus (house mouse)
 Organism Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphati; Muridae; Murinae; Mus.
 Reference 1 (bases 1 to 448)
 Authors Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamamoto,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Built,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Baralov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chochia,C., Corban,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Garboldi,M., Gissi,C., Godzik,A., Gong,J., Grimmond,S., Guscinich,S., Hikokawa,N., Jackson,I.J., Jarvis,B.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kediriski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Mikl,H., Nagashima,T., Numata,K., Okido,T., Pavani,W.J., Pertea,G., Pesoche,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,P.J., Reid,J.J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Seto,M., Shimada,R., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watcana,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yun,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashimoto,K., Ishii,Y., Itoh,M., Kagawa,I., Mizazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E., and Hayashizaki,Y., and vein".
 Note "Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center (GSC), Yokohama Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>
 Title Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Journal Nature
 Medline 2234683
 PubMed 1246851
 Comment Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL:<http://genome.gsc.riken.jp/>

JOURNAL	COMMENT
Hirozane, T., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,	Unpublished (2000)
Myazaki, A., Mura, M., Nakamura, K., Numasaki, R.,	Contact: Robert B. Weiss
Ono, M., Sakai, K., Sakakura, N., Sasaki, D., Sato, K.,	University of Utah Genome Center
Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and	University of Utah
Hayashizaki, Y. Direct Submission	Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
Computational Analysis of Full-length Mouse cDNAs Compared with	84112, USA
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)	Tel: 801 585 5606
Normalization and subtraction of cap-trapper-selected cDNAs to	Fax: 801 585 7177
prepare full-length cDNA libraries for rapid discovery of new	Email: daunn@genetics.utah.edu
genes. Genome Res. 10 (10), 1617-1630 (2000)	Insert Length: 10000 Std Error: 0.00
RIKEN integrated sequence analysis (RISA) System-384-format	Plate: 022 Row: H column: 14
sequencing pipeline with 384 multicapillary sequencer. Genome Res.	Seq primer: CACAGGAAACGCTATGCC
10 (11), 1757-1771 (2000)	Class: plasmid ends
Computer-based methods for the mouse full-length cDNA	High quality sequence stop: 570.
encyclopedia: real-time sequence clustering for construction of a	Location/Qualifiers
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)	1. -448
CDNA library was prepared and sequenced in Mouse Genome	/organism="Mus musculus"
Encyclopedia Project of Genome Exploration Research Group in Riken	/mol_type="mRNA"
Genomic Sciences Center and Genome Science Laboratory in RIKEN.	/db_xref="taxon:10090"
Division of Experimental Animal Research in Riken contributed to	/clone="1920162d12"
please mouse tissues.	/lab_host="B. Coli strain XL10-Gold, T1-resistant, F-
further details.	/clone_lib="Mouse 10kb plasmid UGGCIM library"
Location/Qualifiers	/note="Vector: PWD42nv; purified genomic DNA from the Jackson
1. -570	musculus C57BL/6J (male)" was obtained from the Jackson
/mol_type="genomic DNA"	laboratory Mouse DNA Resource
/strain="C57BL/6J"	(http://www.Jax.org/resources/documents/dnars/). The DNA
/db_xref="taxon:10090"	was hydrodynamically sheared by repeated passage through a
/clone="UUGCIM0222H14"	0.005 inch orifice at constant velocity. The sheared DNA
/sex="Male"	was blunt end-repaired with T4 DNA polymerase and T4
/clone_lib="Mouse 10kb plasmid UGGCIM library"	polynucleotide kinase. Adaptor oligonucleotides were
/note="Vector: PWD42nv; purified genomic DNA from the Jackson	ligated to the blunt ends in high molar excess. The
musculus C57BL/6J (male)" was obtained from the Jackson	adapted DNA was purified and size-selected for a 9.5 to
laboratory Mouse DNA Resource	10.5 kb range using preparative agarose gel
(http://www.Jax.org/resources/documents/dnars/). The DNA	electrophoresis. Vector DNA was prepared from a derivative
was hydrodynamically sheared by repeated passage through a	of PWD42 (gi:47321149b) AF297021, a copy-number
0.005 inch orifice at constant velocity. The sheared DNA	inducible derivative of plasmid R1. The vector was ligated
was blunt end-repaired with T4 DNA polymerase and T4	with adaptors complementary to the insert adaptors and
polynucleotide kinase. Adaptor oligonucleotides were	purified. The sheared, adapted mouse DNA was annealed to
ligated to the blunt ends in high molar excess. The	adapted vector DNA, and transformed into
adapted DNA was purified and size-selected for a 9.5 to	chemically-competent <i>E. coli</i> XL10-Gold (Stratagene) cells
10.5 kb range using preparative agarose gel	and selected for ampicillin resistance."
electrophoresis. Vector DNA was prepared from a derivative	
of PWD42 (gi:47321149b) AF297021, a copy-number	
inducible derivative of plasmid R1. The vector was ligated	
with adaptors complementary to the insert adaptors and	
purified. The sheared, adapted mouse DNA was annealed to	
adapted vector DNA, and transformed into	
chemically-competent <i>E. coli</i> XL10-Gold (Stratagene) cells	
and selected for ampicillin resistance."	
RESULT 8	ORIGIN
Query Match	Best Local Similarity 24.2%; Score 33.6; DB 5; Length 448;
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;	Best Local Similarity 24.2%; Score 33.6; DB 8; Length 570;
DEFINITION	Best Local Similarity 63.8%; Pred. No. 22; Mismatches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
ACCESSION	Matches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
VERSION	Matched sequence:
KEYWORDS	QY 1 GTGACTCTGCTGCGAAGCTAAATAATGCCTGCCTCACCTCTTCAGAAGAGG 60
SOURCE	Db 216 GTGACTTTAGCTAACGTAATTATTCCTCAAATTCCTCATGCTTACTATAAACAGG 157
ORGANISM	QY 61 GTGACTTATTTGCTGGTTTA 80
RESULT 9	Db 156 GAGAATGTTGTTCCCTTA 137
Query Match	Best Local Similarity 24.2%; Score 33.6; DB 8; Length 570;
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;	Best Local Similarity 63.8%; Pred. No. 22; Mismatches 51; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
DEFINITION	Matched sequence:
ACCESSION	BEI31381
KEYWORDS	LOCUS BEI31381
SOURCE	DEFINITION L48-1355T3 Ice plant Lambda Uni-ZAP XR expression library, 48 hours
ORGANISM	NACL treatment Mesembryanthemum crystallinum
RESULT 9	MRNA sequence.
Query Match	Best Local Similarity 59.0 bp mRNA 1 linear EST 20-FEB-2001
Matches 1;保守性 0;差異数 29; 空白数 0; 間隔数 0;	ACCESSION BEI31381
DEFINITION BEI31381	EST. Mesembryanthemum crystallinum (common iceplant)
KEYWORDS BEI31381	Mesembryanthemum crystallinum
SOURCE BEI31381.1 GI:8578744	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
ORGANISM BEI31381.1 GI:8578744	Spermatophyta; Magnoliopsida; eudicots; core eudicots;
RESULT 9	Caryophyllales; Aizoaceae; Mesembryanthemum.
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	
KEYWORDS	
SOURCE	
ORGANISM	
RESULT 9	
Query Match	
DEFINITION	
ACCESSION	</

REFERENCE 1. (bases 1 to 590)

AUTHORS . Cushman, J.C.

TITLE An expressed sequence tag database for the common ice plant, *Mesembryanthemum crystallinum*

JOURNAL Unpublished (1997)

COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR PRIMERS FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 059 row: H column: 01
Seq primer: T3 20mer

HIGH QUALITY SEQUENCE STOP: 605

FEATURES source

Location/Qualifiers

1. . 605
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
/clone="MCR059H01"

/tissue_type="leaf"
/dev_stage="six-week-old"
/clone_lib="ice plant Lambda Uni-Zap XR expression library", 48 hour NaCl treatment prescreened for removal of highly abundant transcripts
/note="vector: Lambda Uni-Zap XR, Bluescript SK-, Site_1: EcorI; Site_2: XbaI"

ORIGIN

Query Match 24.2%; Score 33.6; DB 2; Length 590;

Best Local Similarity 61.4%; Pred. No. 22; Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 44 TTTCCTCAGAAAGAGGCTGACTATTGCTGGTTAACGTGTTATCCCAAAGCAC 103
Db 147 TTTCATCACAGACGGTGAATTCCTAATCCCTAACATCAC 206

Qy 104 CATAATCAGCTAGACTGTTCTATGG 131
Db 207 CTGAAAAAACAGTGGAAGTTGTTTG 234

RESULT 10

LOCUS BM658097

DEFINITION MCR059H01_69977 Ice plant Lambda Uni-Zap XR expression library, 48 hours NaCl treatment prescreened for removal of highly abundant transcripts *Mesembryanthemum crystallinum* cDNA clone MCR059H01_5, mRNA sequence.

ACCESSION CA835622

VERSION CA835622.1

KEYWORDS EST.

SOURCE

ORGANISM *Mesembryanthemum crystallinum* (common iceplant)

REFERENCE 1. (bases 1 to 625)

AUTHORS Cushman, J.C.

TITLE An expressed sequence tag database for the common ice plant, *Mesembryanthemum crystallinum*

JOURNAL Unpublished (1997)

COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR PRIMERS FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 039 row: B column: 06
Seq primer: T3 20mer

HIGH QUALITY SEQUENCE STOP: 625

FEATURES source

Location/Qualifiers

1. . 625
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
/clone="MCS039B06"

Query Match 24.2%; Score 33.6; DB 6; Length 633;
 Best Local Similarity 61.4%; Pred. No. 22;
 Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 44 TTTCCTCAGAAGAGGGTCACTATTGTCGTTATTACTGTTATCCCAAAGCAC 103
 Db 156 TTTCATCAGAACAGCAGTGACTTGTGGTAATCTTACCCATAATCACACATCAC 215

Qy 104 CATAATCAGCTAGACTGTCTTATG 131
 Db 216 CTGAAAAAACAGTGGGAGTTGTTTG 243

ORIGIN

Query Match 24.2%; Score 33.6; DB 6; Length 625;
 Best Local Similarity 61.4%; Pred. No. 22;
 Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 44 TTTCCTCAGAAGAGGGTCACTATTGTCGTTATTACTGTTATCCCAAAGCAC 103
 Db 165 TTTCATCAGAACAGCAGTGACTTGTGGTAATCTTACCCATAATCACACATCAC 224

Qy 104 CATAATCAGCTAGACTGTCTTATG 131
 Db 225 CTGAAAACAGTGGGAGTTGTTTG 252

RESULT 12

LOCUS CAB34105 CAA34105 633 bp mRNA linear EST 12-DEC-2002
 DEFINITION MCS02BH02_153488 Ice plant Lambda Uni-Zap XR expression library, 5 days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase I (2 AM). Mesembryanthemum crystallinum cDNA clone MCS02BH02 5, mRNA sequence.

ACCESSION CAB34105 GI:26561870

VERSION EST.

KEYWORDS Mesembryanthemum crystallinum (common iceplant)

SOURCE Mesembryanthemum crystallinum (common iceplant)

ORGANISM Mesembryanthemum crystallinum (common iceplant)

REFERENCE CUSHMAN, J.C. An expressed sequence tag database for the common ice plant, unpublished (1997).

AUTHORS Cushman, J.C.

JOURNAL Unpublished (1997)

COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR PRIMERS FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 046 row: F column: 12
Seq primer: T3 20mer
High quality sequence stop: 674.
Location/Qualifiers 1. -674

FEATURES source

ORGANISM "Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 /clone="MCS02BH02"
 /tissue_type="leaf"
 /dev_stage="five-week-old"
 /clone="MC046F12"
 /tissue_type="leaf"
 /dev_stage="six-week-old"
 /clone_lib="Ice plant Lambda Uni-Zap XR expression library, 48 hours NaCl treatment prescreened for removal of highly abundant transcripts"
 /note="Vector: Lambda Uni-Zap XR, Bluescript SK-, Site 1: EcoRI; Site 2: XbaI; Library construction was performed according to Stratagene's recommended protocol for the Lambda Uni-ZapXR vector and cDNA synthesis kit."

ORIGIN

Query Match 24.2%; Score 33.6; DB 4; Length 674;
 Best Local Similarity 61.4%; Pred. No. 23;
 Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 44 TTTCCTCAGAAGAGGGTCACTATTGTCGTTATTACTGTTATCCCAAAGCAC 103
 Db 156 TTTCATCAGAACAGCAGTGACTTGTGGTAATCTTACCCATAATCACACATCAC 215

Qy 104 CATAATCAGCTAGACTGTCTTATG 131
 Db 216 CTGAAAAAACAGTGGGAGTTGTTTG 243

RESULT 13

LOCUS BM301497 BM301497 674 bp mRNA linear EST 22-JAN-2002
 DEFINITION MCRO46F12_26036 Ice plant Lambda Uni-Zap XR expression library, 48 hours NaCl treatment prescreened for removal of highly abundant transcripts Mesembryanthemum crystallinum cDNA clone MCRO46F12 5, mRNA sequence.

ACCESSION BM301497

VERSION EST.

KEYWORDS Mesembryanthemum crystallinum (common iceplant)

SOURCE Mesembryanthemum crystallinum (common iceplant)

ORGANISM Mesembryanthemum crystallinum (common iceplant)

REFERENCE CUSHMAN, J.C. An expressed sequence tag database for the common ice plant, unpublished (1997).

AUTHORS CUSHMAN, J.C.

JOURNAL Unpublished (1997)

COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu

PCR PRIMERS FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 046 row: F column: 12
Seq primer: T3 20mer
High quality sequence stop: 674.
Location/Qualifiers 1. -674

FEATURES source

ORGANISM "Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 /clone="MC046F12"
 /tissue_type="leaf"
 /dev_stage="six-week-old"
 /clone_lib="Ice plant Lambda Uni-Zap XR expression library, 48 hours NaCl treatment prescreened for removal of highly abundant transcripts"
 /note="Vector: Lambda Uni-Zap XR, Bluescript SK-, Site 1: EcoRI; Site 2: XbaI; Library construction was performed according to Stratagene's recommended protocol for the Lambda Uni-ZapXR vector and cDNA synthesis kit."

ORIGIN

RESULT 14
 BM301422 LOCUS MW045F12_25886 DEFINITION Ice plant Lambda Uni-Zap XR expression library, 48 hours NaCl treatment prescreened for removal of highly abundant transcripts Mesembryanthemum crystallinum cDNA clone MCRO45F12_5, mRNA sequence.

ACCESSION BM301422
 VERSION BM301422.1 GI:18023797
 KEYWORDS EST.
 SOURCE Mesembryanthemum crystallinum (common iceplant)
 ORGANISM Mesembryanthemum crystallinum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Faboideae; Medicago.

REFERENCE 1 (bases 1 to 746)
 AUTHORS Cushman, J.C.
 TITLE An expressed sequence tag database for the common ice plant, Mesembryanthemum crystallinum
 JOURNAL Unpublished (1997)
 COMMENT Contact : Cushman JC
 Department of Biochemistry
 University of Nevada
 MS300, Reno, NV 89557-0014, USA
 Tel: 775-784-1918
 Fax: 775-784-1650
 Email: jcushman@unr.edu

PCR PRIMERS FORWARD: T3 20mer
 BACKWARD: T7 21mer
 Plate: 045 Row: F column: 12
 Seq primer: T3 20mer

High quality sequence stop: 699.
 Location/Qualifiers 1. 699

FEATURES source
 ORIGIN /organism="Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 /clone="MCRO45F12"
 /tissue_type="leaf"
 /dev_stage="six-week-old"
 /clone_id="Ice plant Lambda Uni-Zap XR expression library, 48 hours NaCl treatment prescreened for removal of highly abundant transcripts"
 /note="vector: Lambda Uni-Zap XR, Bluescript SK-; site_1: ECORI; site_2: XbaI"
 /score: XbaI"

ORIGIN

Query Match Similarity 24.2%; Score 33.6; DB 9; Length 746;
 Best Local Similarity 61.4%; Pred. No. 23; Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
 Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy	Db	Qy	Db
50	517	TCAGAAGAGGGTCACTATTGTCGTGTTAACTGTTATCCCAAAGCACCATT	TGATTAGAGAGTGATGATAGTCATTACATACAGATAATCCCAAAGGAGCACAT
110	577	137	604
CAACTGATAGTGATATCCAAATGACA	CAATGATATCATATCCAAATGACA		

Search completed: August 4, 2005, 01:47:13.
 Job time : 7078 SECs

RESULT 15
 CR303205 LOCUS 746 bp DNA linear GSS 28-FEB-2004 DEFINITION Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago truncatula, genomic survey sequence.

ACCESSION CR303205 DEFINITION
 VERSION CR303205.1 GI:44708225
 KEYWORDS GSS:
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Faboideae; Medicago. (bases 1 to 746)
 Genoscope. Direct Submission Submitted (25-FEB-2004) Genoscope - Centre National de Séquençage : Submitted (25-FEB-2004) Genoscope - Centre National de Séquençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cnrs.fr - Web : www.genoscope.cnrs.fr)
 FEATURES Source
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /cultivar="Jemalong A17"
 /db_xref="taxon:3880"
 /clone_lib="MTEL"
 /note="vector: pIndigoBAC ; site 1: ECORI ; site 2: ECORI ; Debelle F. and Chalhoub B.-Genoscope sequence ID : mtel-21P8FM1"
 /score: XbaI"

REFERENCE AUTHORS TITLE JOURNAL

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Faboideae; Medicago. (bases 1 to 746)
 Genoscope. Direct Submission Submitted (25-FEB-2004) Genoscope - Centre National de Séquençage : Submitted (25-FEB-2004) Genoscope - Centre National de Séquençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cnrs.fr - Web : www.genoscope.cnrs.fr)

FEATURES Source
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /cultivar="Jemalong A17"
 /db_xref="taxon:3880"
 /clone_lib="MTEL"
 /note="vector: pIndigoBAC ; site 1: ECORI ; site 2: ECORI ; Debelle F. and Chalhoub B.-Genoscope sequence ID : mtel-21P8FM1"
 /score: XbaI"

REFERENCE AUTHORS TITLE JOURNAL

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Faboideae; Medicago. (bases 1 to 746)
 Genoscope. Direct Submission Submitted (25-FEB-2004) Genoscope - Centre National de Séquençage : Submitted (25-FEB-2004) Genoscope - Centre National de Séquençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cnrs.fr - Web : www.genoscope.cnrs.fr)

This Page Blank (uspto)

CC linked to a nucleic acid encoding a protein that is able to induce a
 CC protective immune response against an organism in a mammal. The construct
 CC is useful for enhancing expression of a desired protein at mucosal
 CC effector sites. It is useful for delivering a variety of antigenic agents
 CC which can be used to induce a protective immune response against a wide
 range of pathogens such as *Bacillus anthracis*, *Bordetella pertussis*,
 CC *Schistosoma mansoni*, *herpes simplex virus*, and *Mycobacterium tuberculosis*.
 The three promoters (*P(phoP*), *P(pagC*) and *P(ompC)*) are induced at
 CC different stages in the infection process, and hence at different sites
 in the body. This approach allows the induction of different immune
 CC responses which provide protection against pathogens which colonise
 CC different host cell compartments. The *Salmonella* vaccine vector system is
 CC ideally suited to the delivery of many vaccine antigens since the vaccine
 CC delivery mechanism accurately mimics the natural disease, entering the
 XX body via the gut.

Sequence 139 BP; 37 A; 33 C; 23 G; 46 T; 0 U; 0 Other;

Query Match Score 139; DB 4; Length 139;
 Best Local Similarity 100.0%; Pred. No. 6.3e-36; Matches 139; Conservatve 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACTCTGGTCAGCAACTTAATAATGCCTGCCACCTCTTCTAGAMAGGG 60
 Db 1 GTGACTCTGGTCAGCAACTTAATAATGCCTGCCACCTCTTCTAGAMAGGG 60
 Qy 61 GTGACTATTGTCTGGTTATRACTCTTATCCCAAAGCACCATATCACGCTAGAC 120
 Db 61 GTGACTATTGTCTGGTTATRACTCTTATCCCAAAGCACCATATCACGCTAGAC 120

Qy 121 TGTCTTATTGTTAACACA 139
 Db 121 TGTCTTATTGTTAACACA 139

RESULT 2

ID ADG31136 standard; DNA; 2802 BP.

AC ADG31136;

XX DT 26-FEB-2004 (first entry)

XX DE Salmonella typhimurium PphoPQ and phoPQ operon DNA.

XX KW live attenuated derivative; pathogenic Enterobacteriaceae; vaccine;
 KW cross protective immunity; antibacterial; immunostimulant; vaccine;
 KW PhoPQ; promoter; phoPQ operon; ds; gene.

XX OS Salmonella typhimurium.

XX PN WO200396812-A1.

XX PD 27-NOV-2003.

XX PP 15-APR-2003; 2003WO-US011802.

XX PR 15-APR-2002; 2002US-0372616P.

PR 18-APR-2002; 2002US-0373626P.

XX PA (UNIW) UNIV WASHINGTON.

XX PT Curtiss R.;

XX DR WPI; 2004-042484/04; ADG31137, ADG31139.

XX DR P-PSDB; ADG31137, ADG31139.

PT New live attenuated derivative of a pathogenic Enterobacteriaceae species, useful as a vaccine for inducing cross protective immunity against infections caused by various Enterobacteriaceae strains or serotypes.

PT Example 17; FIG 28; 133pp; English.

XX The invention relates to a novel live attenuated derivative of a
 CC pathogenic Enterobacteriaceae species having enhanced ability to induce
 CC cross protective immunity against Enterobacteriaceae. The derivative of
 the invention demonstrates antibacterial and immunostimulant activities
 CC and may be useful as a vaccine for inducing a high level immune response
 and/or cross protective immune response to protect individuals from
 CC infection from a diversity of species or serotypes of bacterial
 pathogens. The current sequence is that of the *Salmonella typhimurium*
 CC *PhoPQ* and *phoPQ* operon DNA of the invention.

XX Sequence 2802 BP; 721 A; 702 C; 724 G; 655 T; 0 U; 0 Other;

Query Match Score 139; DB 12; Length 2802;
 Best Local Similarity 100.0%; Pred. No. 1.5e-35; Matches 139; Conservatve 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACTCTGGTCAGCAACTTAATAATGCCTGCCACCTCTTCTAGAMAGGG 60
 Db 511 GTGACTCTGGTCAGCAACTTAATAATGCCTGCCACCTCTTCTAGAMAGGG 570
 Qy 61 GTGACTATTGTCTGGTTATRACTCTTATCCCAAAGCACCATATCACGCTAGAC 120
 Db 571 GTGACTATTGTCTGGTTATRACTCTTATCCCAAAGCACCATATCACGCTAGAC 630

Qy 121 TGTCTTATTGTTAACACA 139
 Db 631 TGTCTTATTGTTAACACA 649

RESULT 3

ID ABL57265 standard; DNA; 41 BP.

XX AC ABL57265;

XX DT 09-APR-2002 (first entry)

XX DE Escherichia coli DNA 5' to phoA coding region.

XX FH Transcription terminator; trpH; attenuation; vaccine; virucide;
 KW antibacterial; fungicide; antiparasitic; protozoicide; phoA; ds.

XX OS Escherichia coli.

XX FT Key Location/Qualifiers

FT misc_feature 1..19

FT /tag= "a" single-stranded overhang"

FT misc_feature 41

FT /tag= "b" overhang on complementary strand of 4 bases

FT with sequence 5'-GTC-3' "

XX PN WO200230457-A2.

XX PD 18-APR-2002.

XX PR 11-OCT-2001; 2001WO-US031606.

XX PR 12-OCT-2000; 2000US-00689123.

XX (UNIW) UNIV WASHINGTON.

PA (MEGA-) MEGAN HEALTH INC.

PA (CURTIS) CURTISS R.

PA (TING) TINGE S. A.

XX Curtis R. Ting S.A.

XX WPI; 2002-4415/47.

XX PT Composition comprising microbe having attenuating mutation that comprises insertion sequence containing recombinant transcription terminator,

PT useful as vaccine, and for delivering a desired gene product to
 PT individual.

XX

PS Example 5; Fig 11; 91pp; English.

XX The present sequence is that of DNA located 5' to the phoA coding sequence of Escherichia coli. The sequence was used to illustrate an example of the invention relating to deletion of the phoA gene and replacement with a trpA terminator. The phoA mutation was introduced into Salmonella typhimurium SL1344 to produce strain MN-1362. The present invention is based on the discovery that transcription terminators (trp) can be inserted in a bacterial gene in an attenuating strategy which not only attenuates the virulence of the bacteria, but also serves to restrict the effect of the attenuating mutation to the target gene or operon of the bacteria. A claimed vaccine comprises a microorganism having an attenuating mutation in a chromosomal gene, the mutation comprising an insertion sequence which contains a recombinant TT, such as trpA. The microorganism is preferably Salmonella, Shigella or Escherichia, and the TT is preferably inserted in the phoA gene. The vaccine may further comprise a recombinant gene encoding a desired gene product from a virus, bacterium, protozoan, parasite or fungus, or encodes an autoantigen, gamete-specific antigen or an allergen. The attenuated microorganism is also used in a claimed method for delivering a desired gene product to an individual

CC a sequence 41 BP; 10 A; 13 C; 4 G; 14 T; 0 U; 0 Other;

SQ

Query Match 28.8%; Score 40; DB 6; Length 41;
 Best Local Similarity 100.0%; Pred. No. 0.0024; Mismatches 40; Indels 0; Gaps 0;

Oy

11 TCGACGAACTTAATAATGCTCTGCCTTACCCCTTTCTT 50
 Db 1 TCGACGAACTTAATAATGCTCTGCCTACCCCTTTCTT 40

RESULT 4

ID ACR49482 standard; DNA; 1385 BP.

XX

AC ACA49482; DT 19-JUN-2003 (first entry)

XX

DE Prokaryotic essential gene #31139.

XX

KW antisense; ds; prokaryotic essential gene; cell proliferation; KW drug design; gene.

XX OS *Salmonella paratyphi*.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 05-SEP-2001; 2001US-00815242.

PR 25-OCT-2001; 2001US-00948993.

PR 06-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0562699P.

PA (ELIT-) ELITRA PHARM INC.
 XX PT Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GA, Yamamoto R, Forsyth RA, Xu RH; WPI; 2003-028926/02. DR P-PSDB; ABU45612.

XX New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.
 XX

PS Claim 14; SEQ ID NO 37352; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of the nucleic acid inhibitors proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation; or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product; or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation, in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

XX Sequence 1385 BP; 353 A; 372 C; 335 G; 321 T; 0 U; 4 Other;

SQ

Query Match 26.8%; Score 37.2; DB 8; Length 1385;
 Best Local Similarity 92.9%; Pred. No. 0.054; Mismatches 39; Indels 0; Gaps 0;

Oy

1 GTCAGCTGGTGACGACTTAATAATGCTCTGCCTACCC 42
 Db 1344 GTCAGCTGGTGACGACTTAATAATGCTCTGCCTACCC 1385

RESULT 5

ID ADE87477 standard; DNA; 266145 BP.

XX

AC ADE87477; DT 29-JAN-2004 (first entry)

XX

DE Fowlpox virus genome DNA.

XX

KW fowlpox virus; FPV; virucide; tuberculostatic; protozoacide; antipyretic; KW cytostatic; hepatotropic; antibacterial; vaccine; malaria; tuberculosis; KW East Coast fever; avipox virus; influenza; hepatitis; KW human papilloma virus; tumour; leishmaniasis; listeriosis; theileria; gene; ds.

XX OS Fowlpox virus.

XX PN WO2003047617-A2.

XX

PD 12-JUN-2003.

XX 02-DEC-2002; 2002WO-GB005411.

XX 30-NOV-2001; 2001GB-00028733.

PR 30-NOV-2001; 2001US-0334649P.

(ISIS-) ISIS INNOVATION LTD.

XX PA
 XX PI Laidlaw S, Skinner M, Hill A, Gilbert S, Anderson R;
 XX DR WPI; 2003-513700/4B.
 XX PR Treating and/or preventing e.g. malaria or tuberculosis, or eliciting an immune response, comprises administering a priming composition and a boosting composition containing a non-replicating viral vector in either PR order.

XX PS Claim 30; SEQ ID NO 1; 302pp; English.

CC The invention relates to a fowlpox virus (FPV) genome which has modifications in one or more wild-type FPV genes. The invention further relates to a novel method for treating and/or preventing a disease in a subject comprising administering two compositions, each containing a non-replicating viral vector. At least one of the compositions comprises a poxvirus vector derived from a fowlpox virus. The novel compositions have the following activities: virucide, tuberculosic, protozoic, antipyretic, cytostatic, hepatotropic, and antibacterial. The non-replicating viral vector is useful in a vaccine for an animal, particularly a mammal such as a primate, specifically human. The priming or boosting composition, or the kit is useful for manufacturing a medicament for treating and/or preventing a disease which is, or results from, a chronic infection such as malaria, tuberculosis or East Coast fever, or for eliciting a T-cell immune response in a subject. Non-cultured CEF cells are useful for growing an avipox virus, such as fowlpox virus. The method or the vaccine may further be used to treat or prevent influenza, hepatitis, human papilloma virus and other viral infections, malignancies such as tumours, leishmaniasis, listeriosis, and trichinellosis. This polynucleotide sequence represents the DNA of the fowlpox virus genome of the invention.

XX SQ Sequence 266145 BP; 92375 A; 41026 C; 40846 G; 91897 T; 0 U; 1 Other;

Query Match 23.5%; Score 32.6; DB 10; Length 266145; Best Local Similarity 55.9%; Pred. No. 8; Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 18 ACTTAAATATGCCGCGCTCACCCCTTCAGAAGAGCTGACTATTGCTGT 77

Db 21181 AGTCACTTATGAGATATTAAATCTTATGAGCAGGAGAGATCCATGGATG 21240

RESULT 6 ABQ72598/C

ID ABQ72598 standard; cDNA; 772 BP.

AC ABQ72598;

XX DT 03-SEP-2002 (first entry)

XX Human MDDT encoding cDNA SEQ ID NO 150.

XX DE Human MDDT encoding cDNA SEQ ID NO 150.

XX Human; MDDT; disease detection and treatment molecule polynucleotide; proliferative disorder; hepatitis; psoriasis; cancer; AIDS; autoimmune disorder; inflammatory disorder; allergy; multiple sclerosis; rheumatoid arthritis; transgenic; gene therapy; antiarteriosclerotic; heptatropic; antiinflammatory; antiparasitic; cytotoxic; anti-HIV; anti-allergic; antiangiogenic; antiatherosclerotic; antigout; neuroprotective; antirheumatic; antiarrhythmic; gene; SB.

XX OS Homo sapiens.

XX PN WO200240715-A2.

XX PD 23-MAY-2002.

XX PA 06-SEP-2001; 2001WO-US027628.

XX PR 05-SEP-2000; 2000US-0229747P.

XX PR 05-SEP-2000; 2000US-0229748P.

XX PR 05-SEP-2000; 2000US-0229749P.

XX PR 05-SEP-2000; 2000US-0229750P.

XX PR 05-SEP-2000; 2000US-0229751P.

XX PR 05-SEP-2000; 2000US-0230583P.

XX PR 05-SEP-2000; 2000US-0230585P.

XX PR 06-SEP-2000; 2000US-0230584P.

XX PR 06-SEP-2000; 2000US-0230515P.

XX PR 06-SEP-2000; 2000US-0230517P.

XX PR 06-SEP-2000; 2000US-0230518P.

XX PR 06-SEP-2000; 2000US-0230519P.

XX PR 06-SEP-2000; 2000US-0230520P.

XX PR 06-SEP-2000; 2000US-0230597P.

XX PR 06-SEP-2000; 2000US-0230599P.

XX PR 06-SEP-2000; 2000US-0230610P.

XX PR 06-SEP-2000; 2000US-0230865P.

XX PR 06-SEP-2000; 2000US-0230988P.

XX PR 06-SEP-2000; 2000US-0230989P.

XX PR 07-SEP-2000; 2000US-0230951P.

XX PR 07-SEP-2000; 2000US-0231163P.

XX PR 07-SEP-2000; 2000US-0231167P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PT Jackson S, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Wright RJ, Gietzen D, Liu TF, Yap PE; PT Hillman JL, Jones AL, Yu JY, Bradley DL, Rohatgi SD, Harris B, Panzer SR, Flores V, PT Dahl CR, Moniyama MG, Bradley DL, Rohatgi SD, Harris B, Panzer SR, Flores V, Roseberry AM, Gerstel EH, Peralta CH, David MH, Daffo A, Malwaha R, Chen AJ, Chang SC, Au AP, Inman RR; PT Daffo A, Malwaha R, Chen AJ, Chang SC, Au AP, Inman RR; DR WPI; 2002-527544/56.

XX DR P-PSDB, ABP51381.

XX PS Claim 1; Page 390-391; 618pp; English.

CC Novel human disease detection and treatment polypeptide, useful in treatment (Mdr) polypeptide (I) selected from a polypeptide having a sequence selected from 254 sequences (ABP1231-ABP51484) given in the specification, a naturally occurring polypeptide comprising a sequence having at least 90% identity to (I) or a biologically active or immunogenic fragment of (I). (I) is useful for screening a compound for effectiveness as an agonist or antagonist for screening a compound that specifically binds (I) or modulates the activity of (I), and for preparing a polyclonal or monoclonal antibody by hybridoma technology.

CC Nucleic acids (II) (ABQ72449-ABQ72700) encoding (I) are useful for screening a compound for effectiveness in altering expression of a target polynucleotide comprising (I) in a sample or for assessing toxicity of a test compound, detecting Mdr in a sample or for assessing a condition or a disease associated with the expression of Mdr in a biological sample, for detecting (I) in a sample, and for purifying (I) from a sample. A composition comprising (I), an agonist or antagonist is useful for treating a disease or condition associated with decreased or increased expression of functional Mdr. (I) or (II) are useful for diagnosing, treating or preventing disorders associated with aberrant expression of Mdr, where the disorders are selected from a cell proliferative disorder such as arteriosclerosis, cirrhosis, hepatitis, psoriasis, and cancer and an autoimmune/inflammatory disorder such as AIDS, Addison's disease, rheumatoid arthritis. (II) are useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to

CC translocation or inversion among normal, carrier or affected individuals
 CC and as hybridisation probes for mapping naturally occurring genomic
 CC sequences
 XX

SQ Sequence 772 BP; 252 A; 147 C; 156 G; 217 T; 0 U; 0 Other;

Query Match 22.3%; Score 31; DB 6; Length 772;
 Best Local Similarity 57.9%; Pred. No. 5.1; Mismatches 40; Indels 0; Gaps 0;
 Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 PT

Oy 45 TTCTTCAGAAGAGGGGAGATTTGCTTATTAACGTCTTACCCAAAGGCC 104
 Db 549 TTCCAATTAANAAAAGTCCTCCATTAGTTACACAGTTTGTCACAAAGGC 490

RESULT 7

ABQ7265/C ABQ72685 standard; cDNA; 810 BP.
 ID ABQ72685;
 AC ACXX
 DT 03-SEP-2002 (first entry)
 DE Human MDDT encoding cDNA SEQ ID NO 237.

XX Human; MDDT; disease detection and treatment molecule polymucleotide; polymeric fragment of (I). (I) is useful for screening a compound for effectiveness as an agonist or antagonist, for screening a compound that specifically binds (I) or modulates the activity of (I), and for preparing a polyclonal or monoclonal antibody by hybridoma technology.

XX Nucleic Acids (II) (ABQ7249-HBQ7270) encoding (I) are useful for screening a compound for effectiveness in altering expression of a target polynucleotide comprising Oligonucleotides and antibodies are useful for detecting MDDT in sample for assessing toxicity of a test compound, in a diagnostic test for a condition or a disease associated with the expression of MDDT in a biological sample, for detecting (I) in a sample, and for purifying (I) from a sample. A composition comprising (I), an agonist or antagonist is useful for treating a disease or condition associated with decreased or increased expression of functional MDDT. (I) or (II) are useful for diagnosing, treating or preventing disorders associated with aberrant expression of MDDT, where the disorders are selected from a cell proliferative disorder such as arteriosclerosis, cirrhosis, hepatitis, psoriasis, and cancer and an autoimmune/inflammatory disorder such as AIDS, Addison's disease, allergy, anaemia, asthma, atherosclerosis, gout, multiple sclerosis or rheumatoid arthritis. (II) are useful for creating knockout humanized animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation or inversion among normal, carrier or affected individuals and as hybridisation probes for mapping naturally occurring genomic sequences

XX SQ Sequence 810 BP; 262 A; 153 C; 165 G; 230 T; 0 U; 0 Other;

Query Match 22.3%; Score 31; DB 6; Length 810;
 Best Local Similarity 57.9%; Pred. No. 5.2; Mismatches 40; Indels 0; Gaps 0;
 Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Oy 45 TTCTTCAGAAGAGGGGAGATTTGCTTATTAACGTCTTACCCAAAGGCC 104
 Db 549 TTCCAATTAANAAAAGTCCTCCATTAGTTACACAGTTTGTCACAAAGGC 490

RESULT 8

AAR81193/C AAR81193 standard; DNA; 19521 BP.
 ID AAR81193
 AC ACXX
 XX AAR81193;
 DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36005.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytosolic; gene therapy; vaccine; metastasis; ds.

PA (INCY-) INCYTE GENOMICS INC.

PI Jackson S, Lincoln SE, Altus CM, Dufour GE, Chalup MS;
 PI Hillman JL, Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TP, Yap PE;

PI Dahl CR, Momiyama MG, Bradley DL, Rohatgi SD, Harris B;
 PI Roseberry AM, Gerstein EH, Peralta CH, David MH, Panzer SR, Flores V;
 PI Daffo A, Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
 XX

DR P-PSDB; ABP51459.

XX Claim 1; Page 435; 618pp; English.

The invention relates to an isolated human disease detection and treatment (MDDT) polypeptide having a sequence selected from 234 sequences (ABP51211-ABP51484) given in the specification, a naturally occurring polypeptide comprising a sequence having at least 90% identity to (I) or a biologically active or immunogenic fragment of (I). (I) is useful for screening a compound for effectiveness as an agonist or antagonist, for screening a compound that specifically binds (I) or modulates the activity of (I), and for preparing a polyclonal or monoclonal antibody by hybridoma technology.

CC Nucleic Acids (II) (ABQ7249-HBQ7270) encoding (I) are useful for screening a compound for effectiveness in altering expression of a target polynucleotide comprising Oligonucleotides and antibodies are useful for detecting MDDT in sample for assessing toxicity of a test compound, in a diagnostic test for a condition or a disease associated with the expression of MDDT in a biological sample, for detecting (I) in a sample, and for purifying (I) from a sample. A composition comprising (I), an agonist or antagonist is useful for treating a disease or condition associated with decreased or increased expression of functional MDDT. (I) or (II) are useful for diagnosing, treating or preventing disorders associated with aberrant expression of MDDT, where the disorders are selected from a cell proliferative disorder such as arteriosclerosis, cirrhosis, hepatitis, psoriasis, and cancer and an autoimmune/inflammatory disorder such as AIDS, Addison's disease, allergy, anaemia, asthma, atherosclerosis, gout, multiple sclerosis or rheumatoid arthritis. (II) are useful for creating knockout humanized animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation or inversion among normal, carrier or affected individuals and as hybridisation probes for mapping naturally occurring genomic sequences

XX SQ Sequence 810 BP; 262 A; 153 C; 165 G; 230 T; 0 U; 0 Other;

Query Match 22.3%; Score 31; DB 6; Length 810;
 Best Local Similarity 57.9%; Pred. No. 5.2; Mismatches 40; Indels 0; Gaps 0;
 Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Oy 45 TTCTTCAGAAGAGGGGAGATTTGCTTATTAACGTCTTACCCAAAGGCC 104
 Db 549 TTCCAATTAANAAAAGTCCTCCATTAGTTACACAGTTTGTCACAAAGGC 490

XX Oy 105 ATATCACGCTGACTCTTATTTAACCA 139
 Db 489 TTCTGTAGTGTAGAATCTCTCAGTCACAAAGCA 455

XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PR 17-JAN-2001; 2001WO-US001354.
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180638P.
 PR 02-MAR-2000; 2000US-0184664P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0191123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 30-JUL-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-02245213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225737P.
 PR 14-AUG-2000; 2000US-0225738P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225271P.
 PR 14-AUG-2000; 2000US-0225417P.
 PR 14-AUG-2000; 2000US-0225737P.
 PR 14-AUG-2000; 2000US-0225738P.
 PR 14-AUG-2000; 2000US-0225739P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226888P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-02288924P.
 PR 01-SEP-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229505P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234224P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235833P.
 PR 29-SEP-2000; 2000US-0236372P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0236803P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239357P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241810P.
 PR 01-NOV-2000; 2000US-0241617P.
 PR 08-NOV-2000; 2000US-0241784P.
 PR 08-NOV-2000; 2000US-0241785P.
 PR 08-NOV-2000; 2000US-024526P.
 PR 08-NOV-2000; 2000US-024527P.
 PR 08-NOV-2000; 2000US-024527P.
 PR 08-NOV-2000; 2000US-024528P.
 PR 08-NOV-2000; 2000US-024529P.
 PR 08-NOV-2000; 2000US-024532P.
 PR 08-NOV-2000; 2000US-024534P.
 PR 08-NOV-2000; 2000US-024610P.
 PR 08-NOV-2000; 2000US-024611P.
 PR 08-NOV-2000; 2000US-024613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 08-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249220P.
 PR 17-NOV-2000; 2000US-0249221P.
 PR 17-NOV-2000; 2000US-0249222P.
 PR 17-NOV-2000; 2000US-0249223P.
 PR 17-NOV-2000; 2000US-0249224P.
 PR 17-NOV-2000; 2000US-0249225P.
 PR 17-NOV-2000; 2000US-0249226P.
 PR 17-NOV-2000; 2000US-0249227P.
 PR 17-NOV-2000; 2000US-0249228P.
 PR 17-NOV-2000; 2000US-0249229P.
 PR 17-NOV-2000; 2000US-0249230P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250161P.
 PR 01-DEC-2000; 2000US-0250162P.
 PR 01-DEC-2000; 2000US-0250163P.
 PR 08-DEC-2000; 2000US-0251668P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251889P.
 PR 06-DEC-2000; 2000US-0251909P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-48326/52.

PT Nucleic acids encoding human immune/haematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure: SEQ ID NO 36005; 3071PP + Sequence Listing; English.
 XX
 AA AK64951 to AK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM8170 to AAM9121. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AK64703 to AK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AK53492 to AK5950 and AAM82169 represent sequences used in the exemplification of the present invention

SQ Sequence 19521 BP; 5672 A; 3465 C; 3902 G; 6482 T; 0 U; 0 Other; 22.2%; Score 30.8; DB 4; Length 19521; Best Local Similarity 63.5%; Pred. No. 15; Matches 47; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 1 GTGACTCTGGAGCACTTAATAGCCCTGCACCTTGCTTCTAGAGAGG 60
 Db 1311 GAACTCTGGCAACTTACATAACTCTTGCTCAGCTCTGGACCGAGAAAGGG 1252

Qy 61 GTGACTTTTGCT 74

Db 1251 CTAACATTCTAT 1238

RESULT 9
 ADA41637/c
 ID ADA41637 standard; DNA; 19521 BP.
 AC ADA41637;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human secreted protein related DNA.
 XX
 KW Human; secreted protein; cancer; hyperproliferative disorder; rheumatoid arthritis; autoimmune disorder; haematopoietic disorder; anaemia; allergic reaction; asthma; cardiovascular disorder; wound healing; cytosolic; immunosuppressive; nootropic; neuroprotective; antiviral; anti-allergic; hepatotropic; antidiabetic; antiinflammatory; vulnerary; cardiant; gene therapy; db; OS Homo sapiens.

RESULT 10
 ADA5769/c
 ID ADA5769 standard; DNA; 19521 BP.
 AC ADA5769;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DB BAC fragment containing human secreted protein gene #505.
 XX
 KW immuno-suppressive; anti-inflammatory; anti-asthmatic; anti-allergic; KW cytosolic; cerebroprotective; neuroprotective; nootropic; PR cardiovascular; anti-atherosclerotic; gene therapy; KW human secreted protein; immune disorder; inflammation; KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EP1108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PR 18-DEC-2000; 2000EP-00127688.
 XX
 PR 15-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOU) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 P-PSDB; AAG92738.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 2992; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-Lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patient did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 SQ Sequence 2799 BP; 715 A; 731 C; 728 G; 625 T; 0 U; 0 Other;
 Query Match 22.0%; Score 30.6; DB 5; Length 349980;
 Best Local Similarity 53.8%; Pred. No. 39; Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 CC
 Qy 17 AACCTAAATAGCCCTGCCTCACCTCTTCAGAAAGGGTGACTATTGTTGG 76
 Db 271 AACGTCAATGATCGCAGGAAGGCATGCTCATGAGAAACTAACCCATGGTGGG 330
 Qy 77 TTATTAATCTGTATCCCCAACGACCATATACTAACGCTAGCTGTTCTATGTT 133
 Db 331 ATCATCGACAGTCGTCACACCAATTGACCGATCCACGTCACCGATCCACCGATGTCGT 387
 RESULT 13
 AAH6533/C
 ID AAH6533 standard; DNA; 349980 BP.
 XX
 AC AAH6833;
 DT 26-SEP-2001 (first entry)
 XX
 C Glutamic acid synthase fragment SEQ ID NO: 7068.
 XX
 DB Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EP1108790-A2.

XX
 PD 20-JUN-2001.
 XX
 PR 18-DEC-2000; 2000EP-00127688.
 XX
 PR 15-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOU) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 P-PSDB; AAG92738.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Disclosure; SEQ ID NO 7068; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-Lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patient did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 Sequence 349980 BP; 81250 A; 9718 C; 90621 G; 80391 T; 0 U; 0 Other;
 Query Match 22.0%; Score 30.6; DB 5; Length 349980;
 Best Local Similarity 53.8%; Pred. No. 39; Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 CC
 Qy 17 AACCTAAATAGCCCTGCCTCACCTCTTCAGAAAGGGTGACTATTGTTGG 76
 Db 200056 AACGTCAATGATCGCAGGAAGGCATGCTCATGAGAAACTAACCCATGGTGGG 199997
 Qy 77 TTATTAATCTGTATCCCCAACGACCATATACTAACGCTAGCTGTTCTATGTT 133
 Db 199996 ATCATCGACAGTCGTCACACCAATTGACCGATCCACCGATCCACCGATGTCGT 199940
 RESULT 14
 AAH6315/C
 ID AAH6315 standard; DNA; 31241 BP.
 XX
 AC AAH6315;
 DT 12-FEB-2004 (first entry)
 XX
 DB Mycoplasma genitalium gene involved in transport and binding.
 XX
 KW Genetic operating system; nanomachine genome; bioreactor; bioremediation;
 KW therapeutic biomolecule; energy conversion system; processing system;
 KW anabolic; catabolic system; biological film; cosmetic application;
 KW coating; ds.
 XX
 OS Mycoplasma genitalium.
 XX
 PN US2003138777-A1.
 XX
 PD 24-JUL-2003.
 XX
 PF 20-SEP-2001; 2001US-00960858.
 XX


```

; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 37352
; LENGTH: 1385
; TYPE: DNA
; ORGANISM: Salmonella paratyphi A
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (72)..(72)
; OTHER INFORMATION: neg, a, t or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (158)..(158)
; OTHER INFORMATION: neg, a, t or c
; US-10-282-122A-37352

; RESULT 2
; Query Match 26.8%; Score 37.2; DB 17; Length 1385;
; Best Local Similarity 92.9%; Pred. No. 0.074; Mismatches 0; Indels 0; Gaps 0;
; Matches 39; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
; QY 1 GTGACTCTGGTGACGACTTAATAATGCCGCTCACCT 42
; Db 1344 GGAGCTCTGGCGACGAACTTAATAATGCCCTACCC 1385

; RESULT 2
; Sequence 92, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 92
; LENGTH: 100596
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-765-790-92/c

; RESULT 3
; Query Match 23.7%; Score 33; DB 22; Length 100596;
; Best Local Similarity 65.8%; Pred. No. 12; Mismatches 0; Indels 25; Gaps 0;
; Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; QY 18 ACTTAATAATGCCGCTCACCTTTTCTCAGAAGAGGTGACTATTGCTGT 77
; Db 34052 ATTGGATATGCTGTTCTCTTAAACAATATGCTTAAATTGATCTATT 33990
; QY 78 TTATTAACTGTT 90
; Db 33992 TTATTAAAGTT 33980

; RESULT 4
; Sequence 1, Application US/10856118
; Publication No. US20050025747A1
; GENERAL INFORMATION:
; APPLICANT: Laidlow, Stephen
; APPLICANT: Skinner, Mike
; APPLICANT: Hill, Adrian V.S.
; APPLICANT: Gilbert, Sarah C.
; APPLICANT: Anderson, Richard
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: 3742-1000-000
; CURRENT APPLICATION NUMBER: US/10/856,118
; CURRENT FILING DATE: 2004-05-27
; PRIOR APPLICATION NUMBER: PCT/GB02/005411
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: GB0128733.3
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 6/0/334,649
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 266145
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; Query Match 23.7%; Score 33; DB 22; Length 100596;
; Best Local Similarity 65.8%; Pred. No. 12; Mismatches 48; Conservative 0; Indels 25; Gaps 0;
; Matches 18 ACTTAATAATGCCGCTCACCTTTCTTCAGAAGAGGGTGAATTTGTCGCT 77
; Ov

```

; OTHER INFORMATION: Fowlpox virus genome
; US-10-856-118-1.

Query Match 23.5%; Score 32.6; DB 21; Length 266145;
Best Local Similarity 55.9%; Pred. No. 24;
Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 18 ACTTAAATAATGCCCTGCCTCCCTTCAGAACGGGACTATTGCTGTGTT 77
Db 21181 ATGTTACTTATGATATATTATATGACGGAGGAAGATCCATGGTAGTG 21240

Qy 78 TATTAATCTTTAATCCCAGGACCATATCAAGCTGACTGCTTA 128
Db 21241 TTATGATGTTCTGTCATCAAAGGGAATATGATCAATATAATGTTTA 21291

RESULT 5
US-10-363-829-150/c
Sequence 150 Application US/10363829
Publication No. US20040142331A1

; GENERAL INFORMATION:

; APPLICANT: Jackson, Stuart E.; Lincoln, Stephen E.;

; APPLICANT: Altus, Christina M.; Dufour, Gerard E.;

; APPLICANT: Chalup, Michael S.; Jackson, Jennifer L.;

; APPLICANT: Jones, Anissa L.; Yu, Jimmy Y.;

; APPLICANT: Wright, Rachel J.; Gietzen, Darryl;

; APPLICANT: Liu, Tommy F.; Yap, Pierre E.;

; APPLICANT: Dahl, Christopher R.; Moniyyana, Monika G.;

; APPLICANT: Harris, Bernard; Roseberry, Lincoln, Ann M.;

; APPLICANT: Gersin, Jr., Edward H.; Peralta, Careyna H.;

; APPLICANT: Flores, Vincent Z.; Daffo, Abel J.;

; APPLICANT: Marwaha, Rakesh; Chen, Alice J.;

; APPLICANT: Chang, Simon C.; Au, Alan P.;

; APPLICANT: Inman, Rebekah R.

; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT

; FILE REFERENCE: PT-1183 USN

; CURRENT APPLICATION NUMBER: US/10/363, 829

; CURRENT FILING DATE: 2003-03-05

; PRIOR APPLICATION NUMBER: PCT/US01/27628

; PRIOR FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/229, 751

; PRIOR FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: US 60/229, 749

; PRIOR FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: US 60/229, 750

; PRIOR FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/229, 751

; PRIOR FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: US 60/229, 749

; PRIOR FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: US 60/229, 748

; PRIOR FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: US 60/230, 583

; PRIOR FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: US 60/230, 517

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 610

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 597

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 517

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 610

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 597

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 506

; PRIOR APPLICATION NUMBER: US 60/230, 597

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 571

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 610

; PRIOR APPLICATION NUMBER: US 60/230, 597

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 506

; PRIOR APPLICATION NUMBER: US 60/230, 597

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: US 60/230, 571

; PRIOR FILING DATE: 2000-09-06

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PERL Program
; SEQ ID NO 150
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: misc feature
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: LI-814261.1:2000SRP08
; US-10-363-829-237

; Query Match 22.3%; Score 31; DB 19; Length 810;
Best Local Similarity 57.9%; Pred. No. 7.1; Mismatches 40; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 45 TTCTTCAGAGAGGGGACTATTGCGGTTTAACTGTTTACCCAAAGGACC 104
Db 549 TTCCATAAAAGTCTTCATTTGTTATTATTCAGTTGTCACAGTTGTCACAGGCT 490

Qy 105 ATAATCACCTAGCTGTTCTATGTTACACA 139
Db 489 TTCTGTTATGCTAGAATCTCTCAGTGTCAACACA 455

; Query Match 22.3%; Score 31; DB 19; Length 810;
Best Local Similarity 57.9%; Pred. No. 7.1; Mismatches 40; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 45 TTCTTCAGAGAGGGGACTATTGCGGTTTACCCAAAGGACC 104
Db 449 TTCTGTTATGCTAGAATCTCTCAGTGTCAACACA 455

; Query Match 22.3%; Score 31; DB 19; Length 772;

RESULT ⁷
Db 549 TTCCATAAAGAACGCTTCCATTTTACAGTTTGCCAAAGGGCT 490 ; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 2992
; LENGTH: 2799
; Sequence 223, Application US/10486706
; Publication No. US20050071088A1
; GENERAL INFORMATION:
; APPLICANT: LANDFIELD, PHILIP W.
; APPLICANT: BLALOCK, ERIC M.
; APPLICANT: CHEN, KU-YI-CHU
; APPLICANT: FOSTER, THOMAS C.
; TITLE OF INVENTION: GENE EXPRESSION PROFILE BIOMARKERS AND THERAPEUTIC TARGETS FOR
; CURRENT APPLICATION NUMBER: US/10486706
; CURRENT FILING DATE: 2004-02-13
; PRIORITY NUMBER: PCT/US02/25607
; PRIOR APPLICATION NUMBER: US 60/311,343
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 223
; LENGTH: 2337
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-10-486-706-223

Query Match 22.3%; Score 31; DB 21; Length 2337;
Best Local Similarity 56.3%; Pred. No. 11; Matches 58; Conservatve 0; Mismatches 45; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 33 CCCTCACCCTCTTCAGAAAGGGTCACTATTGTCTGGTTATACTTTTAT 92 ; Query Match 22.0%; Score 30.6; DB 9; Length 2799;
; Best Local Similarity 53.8%; Pred. No. 16; Matches 63; Conservatve 0; Mismatches 54; Indels 0; Gaps 0;
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TANISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIORITY NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1

Query Match 22.0%; Score 30.6; DB 9; Length 3309400;
Best Local Similarity 53.8%; Pred. No. 3e+02; Matches 63; Conservatve 0; Mismatches 54; Indels 0; Gaps 0;

Qy 17 AACTTAATATGCCCCCTTCTTCAGAAAGGGTCACTATTGTCTGG 76 ; Query Match 22.0%; Score 30.6; DB 9; Length 2799;
; Best Local Similarity 53.8%; Pred. No. 16; Matches 63; Conservatve 0; Mismatches 54; Indels 0; Gaps 0;
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059

RESULT ⁸
; Sequence 2992, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TANISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIORITY NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1

Query Match 22.0%; Score 30.6; DB 9; Length 3309400;
Best Local Similarity 53.8%; Pred. No. 3e+02; Matches 63; Conservatve 0; Mismatches 54; Indels 0; Gaps 0;

Qy 17 AACTTAATATGCCCCCTTCTTCAGAAAGGGTCACTATTGTCTGG 76 ; Query Match 22.0%; Score 30.6; DB 9; Length 2799;
; Best Local Similarity 53.8%; Pred. No. 16; Matches 63; Conservatve 0; Mismatches 54; Indels 0; Gaps 0;
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059

RESULT ⁹
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OZAKI, AKIO
; APPLICANT: TANISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIORITY NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO: 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-1

Query Match 22.0%; Score 30.6; DB 9; Length 3309400;
Best Local Similarity 53.8%; Pred. No. 3e+02; Matches 63; Conservatve 0; Mismatches 54; Indels 0; Gaps 0;

Qy 17 AACTTAATATGCCCCCTTCTTCAGAAAGGGTCACTATTGTCTGG 76 ; Query Match 22.0%; Score 30.6; DB 9; Length 2799;
; Best Local Similarity 53.8%; Pred. No. 16; Matches 63; Conservatve 0; Mismatches 54; Indels 0; Gaps 0;
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059

RESULT ¹⁰
; Sequence 4, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris

APPLICANT: Burgess, Chris
 APPLICANT: Gannon, Allison
 APPLICANT: Harvey, Jeanne
 APPLICANT: Lechner, John F.
 TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
 FILE REFERENCE: 1657/2032
 CURRENT APPLICATION NUMBER: US10/737,082
 CURRENT FILING DATE: 2003-12-16
 PRIOR APPLICATION NUMBER: US 10/737,082
 NUMBER OF SEQ ID NOS: 300
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 4
 LENGTH: 15649
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-737,082-4

RESULT 11
 US-10-765-790-4/c
 Sequence 4, Application US/10765790
 Publication No. US20050130172A1
 GENERAL INFORMATION:
 APPLICANT: Bayer Healthcare LLC
 APPLICANT: Beard, Chris
 APPLICANT: Burges, Chris
 APPLICANT: Gannon, Allison
 APPLICANT: Harvey, Jeanne
 APPLICANT: Lechner, John F.
 APPLICANT: Li, Zheng
 TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
 FILE REFERENCE: 1657/2035
 CURRENT APPLICATION NUMBER: US/10/765,790
 CURRENT FILING DATE: 2004-11-27
 PRIOR APPLICATION NUMBER: US 10/737,082
 NUMBER OF SEQ ID NOS: 300
 SEQ ID NO 4
 LENGTH: 15649
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-765-790-4

Query Match 21.9%; Score 30.4; DB 22; Length 15649;
 Best Local Similarity 53.3%; Pred. No. 40; Mismatches 56; Indels 0; Gaps 0;
 Matches 64; Conservative 0; MisMatches 56; Indels 0; Gaps 0;

Qy 16 GAACTTAATATGCCCTCCCTACCCCTTTCAGAAGAGGGTCACTATTGCTG 75
 Db 2478 GAACCGAGTATTCATGCCTAGTGTATAGTTAGGACTCAGAACTATTTACTG 2419

Qy 76 GTTTAACTGTATCCCCAACCATATAAGCTAGACTGTCTTATGTTAA 135
 Db 2418 GGTCAGTCTGGTTATAAAGGAATAAGACAGCAGCCAGTCGTTCTTA 2359

RESULT 12
 US-09-960-870-13
 Sequence 12, Application US/09960870
 Publication No. US20030134281A1
 GENERAL INFORMATION:
 APPLICANT: Evans, Glen
 TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
 FILE REFERENCE: P-BA 4738
 CURRENT APPLICATION NUMBER: US/09/960,870
 CURRENT FILING DATE: 2001-09-20
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 13
 LENGTH: 31241
 TYPE: DNA
 ORGANISM: M. genitalium
 US-09-960-870-13

RESULT 13
 US-09-960-858-13/c
 Sequence 13, Application US/09960858
 Publication No. US20030138777A1
 GENERAL INFORMATION:
 APPLICANT: Evans, Glen
 TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
 FILE REFERENCE: P-BA 4974
 CURRENT APPLICATION NUMBER: US/09/960,858
 CURRENT FILING DATE: 2001-09-20
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 13
 LENGTH: 31241
 TYPE: DNA
 ORGANISM: M. genitalium
 US-09-960-858-13

Query Match 21.9%; Score 30.4; DB 10; Length 31241;
 Best Local Similarity 55.8%; Pred. No. 53; Mismatches 46; Indels 0; Gaps 0;
 Matches 58; Conservative 0; MisMatches 46; Indels 0; Gaps 0;

Qy 34 CCTCACCTCTTCTCTAGAAGAGGGTCACTATTGCTCTGTTATAACTGTRATC 93
 Db 25866 CCTCACSCAATTTTAACTACAGTGTTACCACTGTTGTTGCTCTGTC 25807

Qy 94 CCCAACGACCAATAACCCCTAGACTGTTCTTATGTTACA 137
 Db 25806 ATTAAGAAGCTATCAAATCTACCCCTTTTGCTAAA 25763

RESULT 14
 US-10-251-668-13/c
 Sequence 14, Application US/10251668
 Publication No. US20040063097A1
 GENERAL INFORMATION:
 APPLICANT: Evans, Glen
 TITLE OF INVENTION: NANOMACHINE COMPOSITIONS AND METHODS OF
 TITLE OF INVENTION: USE
 FILE REFERENCE: P-BA 5414
 CURRENT APPLICATION NUMBER: US/10/251,668
 CURRENT FILING DATE: 2002-09-20

RESULT 15
 US-09-960-870-13/c

```

PRIORITY APPLICATION NUMBER: US 09/960,607
PRIOR FILING DATE: 2001-09-20
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 13
LENGTH: 31241
TYPE: DNA
ORGANISM: M. genitalium
US-10-251-668-13

RESULT 15
US-10-433-287-79/C
Sequence 79, Application US/10433287
; Publication No. US2004013756A1
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F
; APPLICANT: Liang, Yinghua
TITLE OF INVENTION: Identification of Novel Gene Family Members Expressed by
FILE REFERENCE: 180/132, PCTUS
CURRENT APPLICATION NUMBER: US/10/433,287
CURRENT FILING DATE: 2003-11-12
NUMBER OF SEQ ID NOS: 81

SOFTWARE: PatentIn version 3.2
SEQ ID NO 79
LENGTH: 104644
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: genomic DNA
LOCATION: (1)..(104644)
OTHER INFORMATION: n is an undetermined nucleotide (dATP, dCTP, dGTP, or dTTP)

FEATURE:
NAME/KEY: misc_feature
LOCATION: misc_(3)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc_feature
LOCATION: (39)..(39)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc_feature
LOCATION: (73)..(73)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc_feature
LOCATION: (143)..(143)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc_feature
LOCATION: (231)..(231)
OTHER INFORMATION: n is a, c, g, or t

FEATURE:
NAME/KEY: misc_feature
LOCATION: (242)..(242)
OTHER INFORMATION: n is a, c, g, or t

Query Match 21.9%; Score 30.4; DB 18; Length 31241;
Best Local Similarity 55.8%; Prev. No. 53;
Matches 58; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 34 CCTCACCCCTCTTCAGAAGAGCGGTGACTATTGTCGGTTAACTGTTATC 93
Db 25866 CCTCAGGCAATTTCACAGTGTTACCAAGTGTGAGTTACACACTGGTTTGTC 25807
QY 94 CCCAACGACCATAATCAACCTAGCTGTCTTATGTTACA 137
Db 25806 ATTAAGAAGAACGCTACAATCTATACCCCTTTTGCTAAA 25763

FEATURE: NAME/KEY: misc feature
LOCATION: (11659)..(11659)
OTHER INFORMATION: n is a, c, g, or t
FEATURE: NAME/KEY: misc feature
LOCATION: (11659)..(11659)
OTHER INFORMATION: n is a, c, g, or t
FEATURE: NAME/KEY: MS4A7 initial_coding_region
LOCATION: (17493)..(17639)
FEATURE: NAME/KEY: MS4A7_coding_region
LOCATION: (19439)..(19573)
FEATURE: NAME/KEY: MS4A7_coding_region
LOCATION: (21068)..(21124)
FEATURE: NAME/KEY: MS4A7_coding_region
LOCATION: (23741)..(23947)
FEATURE: NAME/KEY: MS4A7_coding_region
LOCATION: (27037)..(27138)
FEATURE: NAME/KEY: MS4A7_coding_region
LOCATION: (28135)..(28210)
FEATURE: NAME/KEY: misc_feature
LOCATION: (32640)..(32640)
FEATURE: NAME/KEY: MS4A5_initial_coding_region
LOCATION: (28135)..(28210)
FEATURE: NAME/KEY: MS4A5_initial_coding_region
LOCATION: (64028)..(64180)
FEATURE: NAME/KEY: MS4A5_coding_region
LOCATION: (65145)..(65277)
FEATURE: NAME/KEY: MS4A5_initial_coding_region
LOCATION: (66764)..(66820)
FEATURE: NAME/KEY: MS4A5_coding_region
LOCATION: (68118)..(68270)
FEATURE: NAME/KEY: MS4A5_coding_region
LOCATION: (66764)..(66820)
FEATURE: NAME/KEY: MS4A5_coding_region
LOCATION: (68202)..(68219)
FEATURE: NAME/KEY: MS4A5_initial_coding_region
LOCATION: (91488)..(91760)
FEATURE: NAME/KEY: MS4A12_coding_region
LOCATION: (95210)..(95347)
FEATURE: NAME/KEY: MS4A12_coding_region
LOCATION: (96148)..(96204)
FEATURE: NAME/KEY: MS4A12_coding_region
LOCATION: (97865)..(97981)
FEATURE: NAME/KEY: MS4A12_coding_region
LOCATION: (100890)..(101000)
FEATURES: NAME/KEY: MS4A12_coding_region
LOCATION: (101182)..(101283)
FEATURES: NAME/KEY: misc_feature
LOCATION: (104561)..(104625)
FEATURE: OTHER INFORMATION: n is a, c, g, or t
US-10-433-287-79

```

Query Match 21.9%; Score 30.4; DB 19; Length 104644;
Best Local Similarity 53.3%; Pred. No. 89;
Matches 64; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 16 GAACTTAATATAGCCCTGCCTCACCCCTTCTTCAGAAAGGGGTGACTTTCCTG 75
Db 88420 GAACCCAGATTATCATGGCTTGGCTATAGTTATGGACTCAGATACTTTTACTG 88361
Qy 76 GTTTATTAACGTTAACCCAAAGGCCATTATCAACGCGAGCTGTCTATGTTAA 135
Db 88360 GCTCAAGTTCTCGTTTATAAAGGAATTAAAGACAGAGGCCAGCTGATCCTTA 88301

Search completed: August 4, 2005, 01:55:06
Job time : 363 secs

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

Run on: August 3, 2005, 21:47:30 ; Search time 105 Seconds
 (without alignments)
 2166.119 Million cell updates/sec

Title: US-10-070-882A-2

Perfect score: 139

Sequence: 1 gtagactctggcagcaact.....ctgttctttatgttaacaca 139

Scoring table: IDENTITY.NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued Patents NA: *
 1: /cggn2_6/podata/1/ina/5A_COMB.seq: *
 2: /cggn2_6/podata/1/ina/5B_COMB.seq: *
 3: /cggn2_6/podata/1/ina/6A_COMB.seq: *
 4: /cggn2_6/podata/1/ina/6B_COMB.seq: *
 5: /cggn2_6/podata/1/ina/PCTRS_COMB.seq: *
 6: /cggn2_6/podata/1/ina/backfile81.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
C 1	33	23.7	52494	US-09-949-016-16498
C 2	32.2	23.7	862304	US-09-949-016-15273
C 3	31.8	22.9	102304	US-09-949-016-12289
C 4	30	21.6	601	US-09-949-016-81832
C 5	30	21.6	601	US-09-949-016-14837
C 6	30	21.6	14712	US-09-949-016-15840
C 7	30	21.6	22823	US-09-949-016-14211
C 8	30	21.6	211049	US-09-949-016-15770
C 9	29.8	21.4	301	US-09-949-313-293
C 10	29.8	21.4	301	US-09-949-313-293
C 11	29.8	21.4	301	US-09-222-149A-293
C 12	29.8	21.4	301	US-09-159-811-293
C 13	29.8	21.4	301	US-09-636-215-293
C 14	29.8	21.4	301	US-09-635-166A-293
C 15	29.8	21.4	301	US-09-638-489-293
C 16	29.8	21.4	301	US-09-552-616A-293
C 17	29.8	21.4	301	US-09-759-143-293
C 18	29.8	21.4	301	US-09-651-236-293
C 19	29.8	21.4	301	US-09-635-166A-293
C 20	29.8	21.4	2143	US-09-673-395A-82
C 21	29.8	21.4	3174	US-09-489-847-86
C 22	29.2	21.0	16593	US-08-961-527-52
C 23	29.0	21.0	3010	US-08-714-910-71
C 24	29.0	21.0	3010	US-09-265-315-71
C 25	29.0	21.0	3010	US-09-285-311-71
C 26	29.0	21.0	3010	US-09-266-417-71
C 27	29.0	21.0	3010	US-09-528-709-71

ALIGNMENTS

RESULT 1
 US-09-949-016-16498/C

; Sequence 16498, Application US/09949016
 ; GENERAL INFORMATION:
 ; Patent No. 6812339
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1Q01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: Fast-SEQ for Windows Version 4.0
 ; SEQ ID NO 16498
 ; LENGTH: 52494
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(52494)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-16498

Query Match 23.7%; Score 33; DB 4; Length 52494;
 Best Local Similarity 65.8%; Pred. No. 0 66; Mismatches 25; Indels 0; Gaps 0;
 Matches 48; Conservative 0;

Db 11879 ATTGAAATATTGCTGTGTTCTCTTAACTATAATGCACTATT 11820

Qy 78 TTTAACTAGTT 90

Db 11819 TTATTAAAGTTT 11807

RESULT 2
 US-09-949-016-15273/C

; Sequence 15273, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIORITY FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 15273
LENGTH: 86273
TYPE: DNA
ORGANISM: Human
US-09-949-016-15273

Query Match      23.2%; Score 32.2; DB 4; Length 86273;
Best Local Similarity 54.7%; Pred. No. 1.6; Mismatches 53; Indels 0; Gaps 0;
Matches 64; Conservative 0; MisMatch 0;

Qy   23 AATATGGCTGCCAACCCCTTCTAGAAGAGGGTGAATTTGTCGTATT 82
Db   45531 AAATCTGTTGACCAAACTTTCATCTGGAAGGATGAGCAATTCTGTCAGT 45472
Oy   83 AACCTGTTATCCCAAGGACCAATAATCAGCTGAACTGTTATGTTACACA 139
Db   45471 TACATTTATTCATCAATTCAACAGAGGAGAGTTGGATTTAATAATA 45415
RESULT 3
US-09-949-016-12589/c
Sequence 12589, Application US/09949016
Patient No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 12589
LENGTH: 102304
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1:-)(102304)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12589

Query Match      22.9%; Score 31.8; DB 4; Length 102304;
Best Local Similarity 59.3%; Pred. No. 2.3; Mismatches 37; Indels 0; Gaps 0;
Matches 54; Conservative 0; MisMatch 0;

Qy   20 TTAAATATGGCTGCCAACCCCTTCTAGAAGAGGGTGAATTTGTCGTATT 79
Db   6060 TTAACTGCAACTCCACCATCTCTTCTGCCAGAAATATAATATGTCATCTG 6001
Oy   80 ATTACTGTTATCCCAAGGACCAATAATC 110
Db   6000 ATTAGTAACATATGACAGAGCAGGATTC 5970
RESULT 4
US-09-949-016-81832
Query Match      21.6%; Score 30; DB 4; Length 601;
Best Local Similarity 54.5%; Pred. No. 1.1; Mismatches 50; Indels 0; Gaps 0;
Matches 60; Conservative 0; MisMatch 0;

Qy   1 GTOACTCTGGTGCGACTAAATAATGCTGCCTCACCTCTTCTGAAAGAGG 60
Db   154 GAGACTCTGGTGCGACACTAAATGCTGCCTCACCTCTTCTGAGAGG 95
Oy   94 GAGATAATGGTCAACTTTACTTAAAGATTAATAGATGTC 45
RESULT 5
US-09-949-016-144837/c
Sequence 144837, Application US/09949016
Patient No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 144837
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-144837

Query Match      21.6%; Score 30; DB 4; Length 601;
Best Local Similarity 54.5%; Pred. No. 1.1; Mismatches 50; Indels 0; Gaps 0;
Matches 60; Conservative 0; MisMatch 0;

Qy   1 GTGACTCTGGTGCGACACTAAATGCTGCCTCACCTCTTCTGAGAGG 60
Db   154 GAGACTCTGGTGCGACACTAAATGCTGCCTCACCTCTTCTGAGAGG 95
Oy   61 GTGACTCTGGTGCGACACTAAATGCTGCCTCACCTCTTCTGAGAGG 110
Db   94 GAGATAATGGTCAACTTTACTTAAAGATTAATAGATGTC 45

```

RESULT 6

US-09-949-016-15840

; Sequence 15440, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 15840

; LENGTH: 14712

; TYPE: DNA

; ORGANISM: Human

; US-09-949-016-15840

Query Match 21.6%; Score 30; DB 4; Length 14712;
 Best Local Similarity 54.5%; Pred. No. 4,4; Mismatches 0; Indels 0; Gaps 0;
 Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 1 GAGACTCTGGTGAGCAACTTAATATGCCGCGCTCACCCCTTTCTTGTGAAGAGG 60
 Db 3243 GAGACTCTGGCAATCACTTAATCTCTCTGCCTGCTAGTTCCTATCTATAATGGG 3302

Qy 61 GTGACATTTCCTGGTTAACGTTATCCCAGGACCATTC 110
 Db 3303 GAGAATATGCTCACTTTACTTTATAAGATTAATAGATAGTC 3352

; RESULT 7

; Sequence 14121, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-09-03

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 14121

; LENGTH: 22823

; TYPE: DNA

; ORGANISM: Human

; US-09-949-016-14121

; Query Match 21.6%; Score 30; DB 4; Length 22823;

; Best Local Similarity 54.5%; Pred. No. 5,3; Mismatches 0; Indels 0; Gaps 0;

; Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

; Qy 1 GTGACTCTGGTGAGCAACTTAATATGCCGCGCTCACCCCTTTCTTGTGAAGAGG 60
 Db 3243 GAGACTCTGGCAATCACTTAATCTCTCTGCCTGCTAGTTCCTATCTATAATGGG 3302

Qy

61 GTGACTCTGGTGAGCAACTTAATATGCCGCGCTCACCCCTTTCTTGTGAAGAGG 110
 Db 3303 GAGAATATGCTCACTTTACTTTATAAGATTAATAGATAGTC 3352

; US-09-949-016-15770/C

; Sequence 15770, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 15770

; LENGTH: 211049

; TYPE: DNA

; ORGANISM: Human

; US-09-949-016-15770

Query Match 21.6%; Score 30; DB 4; Length 211049;
 Best Local Similarity 61.5%; Pred. No. 14; Mismatches 0; Indels 0; Gaps 0;
 Matches 48; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 13 GAGGAACCTAAATATGCCGCGCTCACCCCTTTCTTGTGAAGAGGGGACTATTGT 72
 Db 164299 GAGGCCCTTACTCTTCTTCATCCTGCCTTATCCTCAATAGCTGAAAGATTCAATT 164240

Qy 73 CTGGCTTATTACTGTAT 90
 Db 164239 ATGGTCTTTATTATT 164222

; RESULT 9

; Sequence 293, Application US/09439313

; Patent No. 6329505

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocke, Susan Louise

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Reed, Steven G.

; APPLICANT: Kalos, Michael

; APPLICANT: Ranger, Gary

; APPLICANT: Retter, Mark

; APPLICANT: Soik, John

; APPLICANT: Day, Craig

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; FILE REFERENCE: 210121.427C9

; CURRENT APPLICATION NUMBER: US/09/439,313

; CURRENT FILING DATE: 1999-11-12

; NUMBER OF SEQ ID NOS: 575

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO: 293

; LENGTH: 301

; TYPE: DNA

; ORGANISM: Homo sapien

; US-09-439-313-293

; Query Match

; 21.4%; Score 29.8; DB 3; Length 301;

Best Local Similarity 56.7%; Pred. No. 0.98; Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 43 CTTTCTCAGAAGGGGACTATTGCTGGTTATRACTGTTCACCA 102
Db 30 CTGTCCTACTGAAAGTCGCTATGCTTGAGTACTCTGATCTGACATC 89

Qy 103 CCATAATCAAGCTAGACGTCTTATGTAACCA 139
Db 90 AATCAATCAATGCCCTAGAGCACTGACTGTTAACCA 126

RESULT 10
US-03-352-616A-293
; Sequence 293, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocke, Susan Louise
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 293
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; US-03-352-616A-293

Best Local Similarity 56.7%; Pred. No. 0.98; Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 43 CTTTCTCAGAAGGGGACTATTGCTGGTTATRACTGTTCACCA 102
Db 30 CTGTCCTACTGAAAGTCGCTATGCTTGAGTACTCTGATCTGACATC 89

Query Match 21.4%; Score 29.8; DB 3; Length 301;
Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 103 CCATAATCAAGCTAGACGTCTTATGTAACCA 139
Db 90 AATCAATCAATGCCCTAGAGCACTGACTGTTAACCA 126

RESULT 11
US-09-232-149A-293
; Sequence 293, Application US/09232149A
; Patent No. 6365611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232,149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 293
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-232-149A-293

Query Match 21.4%; Score 29.8; DB 3; Length 301;
Best Local Similarity 56.7%; Pred. No. 0.98; Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 43 CTTTCTCAGAAGGGGACTATTGCTGGTTATRACTGTTCACCA 102
Db 30 CTGTCCTACTGAAAGTCGCTATGCTTGAGTACTCTGATCTGACATC 89

Qy 103 CCATAATCAAGCTAGACGTCTTATGTAACCA 139
Db 90 AATCAATCAATGCCCTAGAGCACTGACTGTTAACCA 126

RESULT 12
US-09-159-812-293
; Sequence 293, Application US/09159812A
; Patent No. 6613872
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C5
; CURRENT APPLICATION NUMBER: US/09/159,812A
; CURRENT FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 306
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 293
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-159-812-293

Query Match 21.4%; Score 29.8; DB 4; Length 301;
Best Local Similarity 56.7%; Pred. No. 0.98; Matches 55; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 43 CTTTCTCAGAAGGGGACTATTGCTGGTTATRACTGTTCACCA 102
Db 30 CTGTCCTACTGAAAGTCGCTATGCTTGAGTACTCTGATCTGACATC 89

RESULT 13
US-09-636-215-293
; Sequence 293, Application US/09636215
; Patent No. 6320922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocke, Susan L.
; APPLICANT: Jiang, Yuguil
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Rettler, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Barrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Ajun
; APPLICANT: Skeiry, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; LENGTH: 301

; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-636-215-293

Query Match 21.4%; Score 29.8; DB 4; Length 301;
Best Local Similarity 56.7%; Pred. No. 0.98; 0; Mismatches 55; Conservative 0; Indels 0; Gaps 0;

QY 43 CTTTCTCTGAAAGGGTCACTATTTGTTTATCCCAAAGCA 102
Db 30 CTGTCCTCACTGAAAGGTGCTGCTAATCTTGACTCTGACATC 89

QY 103 CCATAATCAAGCTAGACTGTTATTGTTACACA 139
Db 90 AATCATCTAATGCCCTAGAGACTGACTGTACACA 126

RESULT 14

US-09-685-166A-293
Sequence 293, Application US/09685166A
Patent No. 6530305

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Darin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuguil
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.

APPLICANT: Ranger, Gary R.

APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darwick

APPLICANT: Li, Samuel

APPLICANT: Wang, Ajun

APPLICANT: Skeiky, Yasir A. W.

APPLICANT: Hepier, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121-427C21

CURRENT APPLICATION NUMBER: US/09/685,166A

CURRENT FILING DATE: 2000-10-10

NUMBER OF SEQ ID NOS: 898

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO: 293

LENGTH: 301

TYPE: DNA

ORGANISM: Homo sapien

US-09-685-166A-293

Query Match 21.4%; Score 29.8; DB 4; Length 301;
Best Local Similarity 56.7%; Pred. No. 0.98; 0; Mismatches 55; Conservative 0; Indels 0; Gaps 0;

QY 43 CTTTCTCTGAAAGGGTCACTATTTGTTTATCCCAAAGCA 102
Db 30 CTGTCCTCACTGAAAGGTGCTGCTAATCTTGACTCTGACATC 89

QY 103 CCATAATCAAGCTAGACTGTTATTGTTACACA 139
Db 90 AATCATCTAATGCCCTAGAGACTGACTGTACACA 126

RESULT 15

US-09-688-489-293
Sequence 293, Application US/09688489

Patent No. 666377

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer Lynn

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121-427D2
CURRENT APPLICATION NUMBER: US/09/688,489
CURRENT FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 338
SOFTWARE: FastSEQ for Windows Version 3.0
SBO ID NO: 293
LENGTH: 301

; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-688-489-293

Query Match 21.4%; Score 29.8; DB 4; Length 301;
Best Local Similarity 56.7%; Pred. No. 0.98; 0; Mismatches 55; Conservative 0; Indels 0; Gaps 0;

QY 43 CTTTCTCTGAAAGGGTCACTATTTGTTTATCCCAAAGCA 102
Db 30 CTGTCCTCACTGAAAGGTGCTGCTAATCTTGACTCTGACATC 89

QY 103 CCATAATCAAGCTAGACTGTTATTGTTACACA 139
Db 90 AATCATCTAATGCCCTAGAGACTGACTGTACACA 126

Search completed: August 4, 2005, 01:48:59
Job time : 107 secs

This Page Blank (uspto)

GenCore version 5.1.6
copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 3, 2005, 21:27:30 ; Search time 6694 Seconds
(without alignments)
1006.166 Million cell updates/sec

Title: US-10-070-882a-2

Perfect score: 139

Sequence: 1 gtgacttctggcagcaact.....ctgttcttattgttaacaca 139

Scoring table: IDENTITY_NUC

Gappen 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank;*

1: gb_ba;*
2: gb_hg;*
3: gb_in;*
4: gb_om;*
5: gb_ov;*
6: gb_pat;*
7: gb_ph;*
8: gb_pl;*
9: gb_pr;*
10: gb_ro;*
11: gb_sts;*
12: gb_sy;*
13: gb_un;*
14: gb_vl;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT 1
AX100259

DEFINITION Sequence 2 from Patent WO019974.
ACCESSION AX100259
VERSION AX100259.1 GI:13519143

KEYWORDS Salmonella typhimurium
SOURCE
ORGANISM

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

REFERENCE Tibball, R.W. and Bullifent, H.L.
AUTHORS Recombinant microorganisms
TITLE Patent: WO 0119974-A 22-MAR-2001;
JOURNAL The Secretary of State for Defence (GB)

FEATURES Source
1. 139
/organism="Salmonella typhimurium"
/mol_type="unassigned DNA"
/db_xref="taxon:602"

ORIGIN

Query Match 100.0%; Score 139; DB 6; Length 139;
Best Local Similarity 100.0%; Pred. No. 7.3e-29;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 60
Dy 1 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 60
Qy 61 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 120
Dy 61 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 120

Query Match 100.0%; Score 139; DB 6; Length 139;
Best Local Similarity 100.0%; Pred. No. 7.3e-29;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 60
Dy 1 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 60
Qy 61 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 120
Dy 61 GCGACTTCTGGCAGCACTTAATATGCGCTTCACTCCCTTTCTTCGAAGAGG 120
Qy 121 TGTCTTATGTAAACA 139
Dy 121 TGTCTTATGTAAACA 139

RESULT 2
STYHOPA

Locus	STY27220	STY27220	990 bp	DNA	linear	BCT 26-APR-1993	Title	Salmonella enterica serovar Typhimurium response involved in attenuation of pathogen intracellular proliferation
Definition	S.typhimurium phop gene encoding Phop virulence protein	S.typhimurium phop gene encoding Phop virulence protein	attenuation of pathogen intracellular proliferation	JOURNAL	Infect. Immun.	69 (10), 6463-6474 (2001)	JOURNAL	Infect. Immun.
Accession	cds	cds	6474	MEDLINE	21437654	(2001)	PUBMED	1153591
Version	M22421	M22421.1	GI:154263	REFERENCE	2	(bases 1 to 2459)	AUTHORS	Garcia-del Portillo,F.
Keywords	phop gene; virulence protein.	phop gene; virulence protein.	virulence protein.	ARTICLE	Direct Submission		ARTICLE	
Reference				COMMENT	Submitted (16-FEB-2000) Garcia-del Portillo,F., Centro de Biología Molecular, 'Severo Ochoa', Campus de Cantoblanco, 28049 Madrid, SPAIN		COMMENT	
Authors				FEATURES	location/Qualifiers		FEATURES	
Title				source	1. . 2459	/organism="Salmonella typhimurium"	source	1. . 2459
Journal				RBS	/db_xref="taxon:602"	/mol_type="genomic DNA"	JOURNAL	/organism="Salmonella typhimurium"
Medline				CDS	140. . 142	/strains="SL134"	JOURNAL	/mol_type="genomic DNA"
Pubmed					151. . 825	/db_xref="taxon:602"		/strains="SL134"
Comment					/note="Phop protein"	/db_xref="taxon:602"		
Original source text:	S. typhimurium (strain 14028) DNA, clone PR35381.	Original source text:	S. typhimurium (strain 14028) DNA, clone PR35381.	gene	240. . 914	/function="virulence transcrip-	gene	240. . 914
Drift entry and computer-readable sequence [1] kindly submitted by F.Heffron, 01-JUN-1989.	Location/Qualifiers	/location/Qualifiers	Location/Qualifiers	gene	/gene="phop"	tional regulatory system"	gene	/gene="phop"
FEATURES				CDS	240. . 914	/note="member of a two-component regulatory system"	FEATURES	
Source					/codon_start=1	/codon_start=1	source	
					/transl_table=11	/transl_table=11	JOURNAL	
					/product="Phop protein"	/product="Phop protein"	JOURNAL	
					/protein_id="CAB7592.1"	/protein_id="CAB7592.1"		
					/db_xref="GI:700769"	/db_xref="GOA:P14146"		
					/db_xref="GOA:P14146"	/db_xref="UniProt:Swiss-Prot: P14146"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/translation="MMLVLLVEDNALLRLHLKVQIQLQDSGHQVAEADAREADYLNHEH		
					/transl_table=11	LPPDIAVDPGLPDEGSLIRRKRSRSDSVLPVLVLTARSCWQDDKVEVLSVSSGADDYVTK		
					/product="Phop protein"	PF11KEWVARMOALMRNSGLASQVNINPPFOVDLSRERBLSNVEEVILTAFTYTIME		
					/protein_id="CAB7592.1"	TURNGKVVSKSISLMLQYDAAEADRESHTIDVLMGRLRKIQAQYPHDVITVRQGQ		
					/db_xref="GI:700770"	YLFEIR"		
					/db_xref="GOA:P14147"	914. . 2377		
					/db_xref="UniProt:Swiss-Prot: P14147"	914. . 2377		
					/transl_table=11	/gene="phoQ"		
					/product="PhoQ protein"	/gene="phoQ"		
					/protein_id="CAB7592.1"	/gene="phoQ"		
					/db_xref="GI:700770"	/gene="phoQ"		
					/db_xref="GOA:P14147"	/function="membrane sensor protein"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/function="membrane sensor protein"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P14147"	/note="member of a two-component regulatory system"		
					/db_xref="UniProt:Swiss-Prot: P14147"	/note="member of a two-component regulatory system"		
					/transl_table=11	/note="member of a two-component regulatory system"		
					/product="PhoQ protein"	/note="member of a two-component regulatory system"		
					/protein_id="CAB7592.1"	/note="member of a two-component regulatory system"		
					/db_xref="GI:700770"	/note="member of a two-component regulatory system"		
					/db_xref="GOA:P1414			

RESULT 4	AL627269/c	LOCUS	Al627269	DEFINITION	Salmonella enterica serovar Typhi; (Salmonella typhi), strain Cn18, complete chromosome; segment 5/20.	ACCESSION	Al627269 AL53382	VERSION	Al627269.1 GI:16502231		
ORGANISM	Salmonella enterica subsp. enterica serovar Typhi	TITLE	Salmonella enterica subsp. enterica serovar Typhi	AUTHORS	Salmonealla enterica subsp. enterica serovar Typhi Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.	KEYWORDS	LSDLVYKITALA" QDGGGQPLVEMTDLNSRNQVQASRLIEPLRKRYYDKRQKOMIALEOLKGLEN	REFERENCE	532..-1332		
JOURNAL	Nature 413 (6858), 848-852 (2001)	COMMENT	1 (bases 1 to 25050) Parthill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T.G., Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connerton, P., Crook, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T.S., Leather, S., Moule, S., O'Gaora, P., Party, C., Quail, M., Rutherford, M., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G.	FEATURES	Complete genome sequence of a multiple drug resistant <i>Salmonella</i> enterica serovar Typhi CT18	REFERENCE	PRD_ECOLI (336 aa), 95% identity in 336 aa overlap"	ORGANISM	/gene="STY1078"		
GENE	2 (bases 1 to 254050) Parthill, J.	CDSS	Direct Submission	JOURNAL	Submitted (25-OCT-2001) Submitted on behalf of the <i>Salmonella</i> sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK	COMMENT	Notes: Details of <i>S. typhi</i> sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/S_typhi/).	FEATURES	/db_xref="UniProt:Q8ZTS9" /translation="MYYFVKRKFQDFPERRARFTFOOLRRITGPTEALVROKPT KPTCAGLTFKPNFLGLAAGIDKOCEDGALGAGGDDCQEGTTPRPPGNDPRLFR LVDEGLINNRMGFNNMLVNLVNUVKKHFGDGLGNGINGENKNKTPVENGDDLTICME KVAYAGYTAINTSSPNTNPNGPLRTQYGDADDLITAKNKNDQLOVIIHHKYVPAVKI APDCCBBLIQVADSLRLLNRDVAWITTSLSVQEGNCQGGSPRQLQKST EIRRSLSLBKQQLPPIVGVGIBSVIAKANGATLWQIVSGFPFVKGPPLIKEIKV HI"	GENE	/gene="STY1078"
GENE	3136..3990	misc_feature		GENE	/note="fam match to entry PF01433 Peptidase_M1, Peptidase_N, Pyrd"	CDSS	Dihydroorotate dehydrogenase, score 563.50, E-value 1.4e-165"	GENE	/note="fam match to entry PF01180 DHODHase, score 563.50, E-value 3.238..3.297		
GENE	3238..3297	misc_feature		GENE	/note="STY1079"	CDSS	/gene="STY1080" /note="PF00911 Dihydroorotate dehydrogenase signature 1"	GENE	/note="STY1079"		
GENE	3874..3936	misc_feature		GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."	CDSS	/gene="STY1080" /note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site." /codon_start=1	GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."		
GENE	4174..4719	misc_feature		GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."	CDSS	/transl_table=11 /product="conserved hypothetical protein"	GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."		
GENE	4174..4719	misc_feature		GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."	CDSS	/product="conserved hypothetical protein"	GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."		
GENE	4719..5825	misc_feature		GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."	CDSS	/transl_table=11 /product="conserved hypothetical protein"	GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."		
GENE	5825..5825	misc_feature		GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."	CDSS	/product="conserved hypothetical protein"	GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."		
GENE	5825..5825	misc_feature		GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."	CDSS	/product="conserved hypothetical protein"	GENE	/note="Orthologue of <i>E. coli</i> YCBW_ECOLI, Fasta hit to YCBW_ECOLI (192 aa), 90% identity in 180 aa overlap. Note discrepancy in position of translational start site."		

QY 121 TGTCTTATGTTAACACA 139
 Db 175878 TGTCTTATGTTAACACA 175860

misc_feature

gene

CDS

/db_xref="GOA:Q8ZTS7"
 /db_xref="UniProt:TREMBI:Q8ZTS7"
 /translation="MATSRLRFIPVKSURGIGHALADISGLAFDRIMVTFESDST
 FTRKQPQNRFFSPFLADGHITAPDSSALIRFTDFPQDFTWGNHITARVA
 PTANQWISGFSDVQLCNGPDLTRRYKHNVPIGADPGYLLNEASIRLQQ
 RCPGGVOMSOFRPLVSVGAAMEREDSWKVRGIDWVKUCCFTTSPFKK
 KHPGEPLATIQLARTAQDNGDVGQNLJARNSGVIRKGDEELATPAKAGATT
 LDDSVTPERKHDGAVTIDMQGQTCGNNNQVLLEQLENQGIRIPYSRCAGICCCRIR
 complement (-737..-4955)

/gene="STY1082"

/note="PFam match to entry PRO011 fer2, 2FE-2S
 iron-sulfur cluster binding domains, score 55.90, E-value
 8.5E-13"
 5924..8032

/gene="STY1082"

5924..8032

/gene="STY1082"

/note="Similar to *Haemophilus influenzae* hypothetical
 protein H1015 H1016 SW_YCB_HAEIN (P44524; P43945) Fasta
 scores: E0: 0, 60.3% id in 710 aa
 Orthologue of *E. coli* YCB1_ECOLI; Fasta hit to YCB1_ECOLI
 (702 aa), 92% identity in 702 aa overlap"
 /codon_start=1

/transl_table=11

/product="conserved hypothetical protein"
 /protein_id="RAD08187.1"

/db_xref="GI:16502336"
 /db_xref="GOA:Q8ZTS6"
 /db_xref="UniProt:TREMBI:Q8ZTS6"

/translation="MNLSEASTAGLEELIKTEBEKGAVGCOVQVGHFQGDTRLI
 YQSMWSRLASRSLTDPGSKVYDIFQVQINWTFENPDTFAVHFSGNDLII
 LRGDRDTGLAPIKETLAIVMSQGWQPTPLDMCGSGTLLIEAMWATPAPGL
 HRRGNGFSGIAQHAIWQEVKAQTRAKGLAKISSHYGSDDARVTERARNAR
 RAGIGELITPEVABRNKLNGSLGRHLIQADCLGKHLRANEBOFLIDPPTSNSKR
 NOFGCWNLSLFSASPDLSLQLADKORKAOKGRLDCYOKNTHAFTADSRAVTA
 EDYANRURKRNKKERKWARKEGECIYKLTADDEBKVINAVDRCGDAVWQYAPPKV
 DAKQRKQRLEFDIATLTSVIGIPKLVTDKPKRQKNGYQKNSKBRFLEYSEYNA
 RIWMLNTDYLDTGLFDLHTRMIGEMSGKDKDPLNFESTGVSYHAGLGARNTT
 VDMSMTYLEWABRNKLNGSLGRHLIQADCLGKHLRANEBOFLIDPPTSNSKR
 MEESDPOVORHVALMKDKRKGTTFSNNKRGFRNLDLEGIAELGITAQETIOTKT
 LSDPFPARNRHOHNWLRAA"
 5927..7060

/gene="STY1082"

/note="PFam match to entry PF01170 UPF0020, score 781.30,
 Uncharacterized protein family UPF0020

misc_feature

gene

CDS

6506..6541

/gene="STY1082"

/note="PS01261 Uncharacterized protein family UPF0020
 signature"

8045..9952

/gene="STY1082"

8045..9952

/gene="STY1083"

/note="PFam hit to YBRT_ECOLI (530 aa), 32% identity in
 526 aa overlap
 Fast hit to YOJK_ECOLI (554 aa), 34% identity in 524 aa
 overlap

Query Match 96.5%; Score 134.2; DB 1; Length 251050;
 Best Local Similarity 97.8%; Pred. No. 4; 7e-28;
 Matches 136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCGACTCTGGTCACGAACTTAATAATGCCAGCCCTTCACTCTTCTGAGAAAGGG 60
 /gene="STY1083"
 /note="PFam hit to YBRT_ECOLI (530 aa), 32% identity in
 526 aa overlap
 Fast hit to YOJK_ECOLI (554 aa), 34% identity in 524 aa
 overlap

gene

CDS

1382..2173

/locus_tag="t1460"
 /note="corresponds to STY1521 from Accession AL513382:
 Salmonella typhi Ctl8"
 /codon_start=1

/transl_table=11

/product="putative regulatory protein"
 /protein_id="AA069100.1"
 /db_xref="GI:9137538"
 /translation="METCGCFSLSTCVINYDHQRQDMNPFPKIPWVWIGRMVIP
 EVRERLYKANTILKUDNIPFPTLPPERELAKOGVSPASLRELALVIEISG
 WIVQSGNGTVISDKHLASDYYTIEELISRELDSHCARLAQDNVWQIBAY
 HRMDQAINDNVHGRPSLKDPLAISEASRNRLFDMSRMLWBQRTINIPYAGLBDQS

GDRHNLNLNKKQHKAIVDAMRQSHADSAYEGSLHEHLRYVRKIVCG"

gene complement (2232..3257)

/locus_tag="t1461" complement (2232..3257)

/note="corresponds to STY1520 from Accession AL513382: Salmonella typhi CT18"

CDS /codon_start=1

/transl_table=11

/product="putative alcohol dehydrogenase"

/protein_id="AA069101_1"

/db_xref="GI:29137539"

/translation="MKVKAIFNKGPEGEVETRWWYPHKKEENEVILVKVDAIGICGSDIG

AFRGTPNPLMVPYPRIGHEVYGVIOEGTMDPDKKGDRVIVDPIYTCGHYCPSVGR

TNCCENLNNVJGVHDGAWQEVNTIPKDPVNSMAPLALIPLHAIHQADQ

VKAEBYVATIGAGIGMQLAISVHAKPFLIIVEERLYACKLGIVPVVINAADQ

VDAIKNTHGTGMOWVITASGASAIRWDLASFAGRIISPCTWPKPOTSPLPNLT

VEEDLDRGSRTSAGBDFEALMLSTLEINQFDVSVKVNLDIDFDDAVSKLDYPERYL

KINAVPH"

complement (3232..4518)

/locus_tag="t1462" complement (3232..4518)

/locus_tag="t1462"

/note="corresponds to STY1519 from Accession AL513382: Salmonella typhi CT18"

CDS /codon_start=1

/transl_table=11

/product="membrane transport protein"

/protein_id="AA069102_1"

/db_xref="GI:29137540"

/translation="EMGKTMGWLFLVPLRPTVMYMDMSALSTAPLIBKELGFFNA

EVRLIGMAGEGVSAAGKKTINWNWRSKESIAIGFSASGSPLGAVSGVIGVGLAISL

GWRPAGCITITLPGVNLWVPSUDSKPTKMSKIDFMRHEDVLSDDORARATP

SIGKYMOKPMWATLAFAESYYLFFELWTPSTLNHSALHDLIKEISATVWIVIG

ALGMVJGGVSVDVITRIGMALLSLIGVCLAGAVCAVSJSTGTSALIMSV

SLLFLYLUTGPYAVDQVHDKYQGSVGAAGMHGQNSIGRIGPLVTGFIVQFGKD

YAFVLAGAIIATSVSLVFVFKSKOFKANESQSCVH"

complement (4695..6251)

/gene="pgaa"

/locus_tag="t1463"

/complement (4695..6251)

/product="hypothetical protein"

/protein_id="AA069103_1"

/db_xref="GI:29137541"

/translation="NEKRYIWKLMWLAFLALVURKRGESMMKKLYLVVVLFFPSTKVTLL

HNNIFCRANSFEDLSHVLPDREQISGTPKTYTJSTPLAQVUVVRYHESLQHSP

DDMWVTRQWHRNDVIVDITPETAKEHHLVUVNGINTERGQIPLKPAEVDFQMLNPLNT

RDTNTTVISVSDISVQYLTRODDKKELEDEBSVSEWALMPEAAPSQRELMPLNMT

ALISQANLAKKELQTMINSITGIGSKRGTWMSIADPVAEIVPRAIDLIDA

SLEHLYQSYGGENWPLMVPYQQGIDEKISKPTFTLQRQIDPRLNTYQPLAIP

KYTNASGDDFPPVDPYKSYKCVKSYKCVKSYKCVKSYKCVKSYKCVKSYKCV

KTKTROLGLIHLHLLTVSEFSEPKVURWMPANHARDERYAGGYGRYQHPTIDIPAN

NKISTITNEPKVWEATYIEATIFNDGIVATSQVYITPDEKIPQTAAPSVNACQILPG

RGLIGENDSPD"

complement (6321..7562)

/locus_tag="t1464" complement (6321..7562)

/locus_tag="t1464"

/note="corresponds to STY1517 from Accession AL513382: Salmonella typhi CT18"

CDS /codon_start=1

/transl_table=11

/product="putative multidrug efflux protein"

/protein_id="AA069104_1"

/db_xref="GI:29137542"

/translation="MNTNVYENTDSETITPLNKKRILPVFLVUGLYAASTAVMSVLP

FYLRMGSSPLIGIILTEAFSQCAPLIGHSDRVGRKRILVTLAAISLIL

ANAOCLTFLPLARTLTFGSGASAAVYIADCTHVRURQAGILTGIGLIGIIGA

GVSQSLSRSLIGGAPIYAFTLIGAULVAGLQKDPSTSRTTDKIASFRCATLKM

VLRLITLICHFFAGMNSQSPVFLSDTTWNLGPGLGKPSLQKSYLIMAGVUNIVFO

LFLIGWWSQYFSERKLILIFALCTGTGUTTAGIATTIVLIFAVCISIADALAKPTY

LVTAHYIYLSKSLIJAOKBT"

complement (7552..8714)

/locus_tag="t1465"

/note="corresponds to STY1515 from Accession AL513382: Salmonella typhi CT18"

CDS /pseudo

/locus_tag="t1466"

8759..9247

8759..9247

/locus_tag="t1466"

/note="corresponds to STY1514 from Accession AL513382: Salmonella typhi CT18"

CDS /codon_start=1

/transl_table=11

/product="putative regulatory protein"

/protein_id="AA069105_1"

/db_xref="GI:29137543"

/translation="MSENKNVQDHTISBOMKALHGALTRWVSALNPRBNDKBLIAEG

IQLRHALFLLISLERLGLGIVVSLAERAGRDYTVSROWAKLBGLUIRQHWRID

RIRAVVISRTGKAMTERIDAREQMGNTVVKWDWPODKDIDFVRLMOKFADAMSDASM

PE"

9429..10508

9429..10508

/locus_tag="t1467"

/note="corresponds to STY1513 from Accession AL513382: Salmonella typhi CT18"

CDS /codon_start=1

/transl_table=11

/product="putative isomerase"

/protein_id="AA069106_1"

/db_xref="GI:29137544"

/translation="MLTEKDYFRITDQMTIPLRPHWIANDSYRECKCMVLYNRSKBI

HKVDESKVTDIKEGDVICENDSTINHMRICKRRRLKIVLEGFLINNRVLSL

LKERIANDVYLVNPESKIKLQRGSESQYRAVENBALICLASYLASHGERIDEV

MLSNKIQTCVCEVNTINEYEYEQTAQADINTAKQONGGRIAVGTTVRCLSESAYS

REHNCUKASSGWTAIYTHQQLQYKQVLYDQEGOC" SEDSOCOEMFGDCMIIQDQEGOC"

complement (11372..12544)

/locus_tag="t1468"

/complement (11372..12544)

/locus_tag="t1468"

/note="corresponds to STY1509 from Accession AL513382: Salmonella typhi CT18"

CDS /codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AA069107_1"

/db_xref="GI:29137545"

/translation="MENANASGDDFPPVDPYKSYKCVKSYKCVKSYKCVKSYKCV

VILNKWFDTMHAALITTEVRASKSYKCVKSYKCVKSYKCVKSYKCVKSYKCV

PPSLSHLSLSPSLPDNSYMRRLVLEVLNLRPAIPNTIDDYPLVPELD

EDKRNHEOCGAAFSOLFEDKGRNSETFOKLINPTEQD

NFNFKTNEPKVWEATYIEATIFNDGIVATSQVYITPDEKIPQTAAPSVNACQILPG

RGLIGENDSPD"

complement (6321..7562)

/locus_tag="t1464" complement (6321..7562)

/locus_tag="t1464"

/note="corresponds to STY1517 from Accession AL513382: Salmonella typhi CT18"

Query Match
Best Local Similarity 97.8%; Pred. No. 4; Se-28;
Matches 136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GTGACTCTGGTCAGGACTTAATAAGCCCTCCACCTCTTCTCAGAACAGG 60

Db 250187 GTGAATCTCTGCTGAGCTTAATAAGCTGCGCTACCTCTTCTCAGAACAGG 250246

OY 61 GTGACTCTTGTCGGTTAACTGTTATGCCAAAGCACCATATCAAGCTAGAC 120

Db 250247 GTGACTCTTGTCGGTTAACTGTTATGCCAAAGCACCATATCAAGCTAGAC 250306

RESULT 6		RESULT 7	
QY	Db	QY	Db
TGTTCTTATGTGTAACACA 139	250307 TGTCTTATGTGTAACACA 250325	GTAAACACA 139	121 GTTAACACA 129
STYPHOPQ 2190 bp DNA linear BCT 26-APR-1993	STYPHOPQ 24578 bp DNA linear BCT 23-APR-2003		
LOCUS S.typhimurium phop protein and membrane protein phoQ gene, complete cds.	LOCUS Salmonella typhimurium LT2, section 58 of 220 of the complete genome.		
DEFINITION	DEFINITION AE008754 AE008754		
ACCESSION M24424	ACCESSION AE008754.1 GT:16419750		
VERSION M24424.1 GI:154265	VERSION AE008754.1		
KEYWORDS membrane protein; phop gene; phoQ protein.	KEYWORDS Salmonella typhimurium LT2		
SOURCE Salmonella typhimurium	SOURCE Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.		
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.	ORGANISM		
REFERENCE 1 (bases 1 to 2190)	REFERENCE 1 (bases 1 to 24578)		
AUTHORS Miller, J., Kukar, A.M. and Mekalanos, J.J.	AUTHORS McClelland, M., Sanderson, R.E., Spieth, J., Clifton, S.W., Latresille, P., Courtney, L., Porwollik, S., Ali, J., Dantes, M., Du, F., Hou, S., Leyman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, S., Sun, H., Florea, L., Miller, W., Stoecking, T., Nhan, M., Waterston, R. and Wilson, R.K.		
TITLE typhimurium virulence	TITLE Complete genome sequence of <i>Salmonella enterica</i> serovar Typhimurium LT2		
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (13), 5054-5058 (1989)	JOURNAL Nature 413 (6858), 852-856 (2001)		
PUBMED 89296942	PUBMED 2153948		
COMMENT Original source text: S.typhimurium (strain LT2) DNA Draft entry and printed copy of sequence for [1] kindly provided by S.I. Miller, 01-MAY-1989.	COMMENT Location/Qualifiers 1. 2190		
FEATURES source	FEATURES source		
CDS	CDS		
/organism="Salmonella typhimurium"	/organism="Salmonella typhimurium"		
/mol_type="genomic DNA"	/mol_type="genomic DNA"		
/db_xref="taxon:602"	/db_xref="taxon:602"		
141..815	141..815		
/note="phop protein"	/note="phop protein"		
/codon_start=1	/codon_start=1		
/transl_table=11	/transl_table=11		
/protein_ids="AAA27188..1"	/protein_ids="AAA27188..1"		
/db_xref="GI:154266"	/db_xref="GI:154266"		
/translation=MNRVQVEDNDALLRRHLYKVQLODGSQHQAADAREADYYNEH	/translation=MNRVQVEDNDALLRRHLYKVQLODGSQHQAADAREADYYNEH		
LPIIAIWVGLDDEGGSLIRPWRSSVSPVLMVTARBGKDKVETLSSGADDYVTK	LPIIAIWVGLDDEGGSLIRPWRSSVSPVLMVTARBGKDKVETLSSGADDYVTK		
PFHIEWMARMQMLRNGNGLASQVINIPPPFOVDSLRSRELSENEVILKLTAFYTIME	PFHIEWMARMQMLRNGNGLASQVINIPPPFOVDSLRSRELSENEVILKLTAFYTIME		
TYRNNGAVVKSKDLSMLQLYPDABLRBESHTIDVLMGRLRKKIQAQYPHDVITTVRGOG	TYRNNGAVVKSKDLSMLQLYPDABLRBESHTIDVLMGRLRKKIQAQYPHDVITTVRGOG		
VIFLR" 815..2170	VIFLR" 815..2170		
/note="membrane protein phoQ"	/note="membrane protein phoQ"		
/codon_start=1	/codon_start=1		
/transl_table=11	/transl_table=11		
/protein_ids="AAA27189..1"	/protein_ids="AAA27189..1"		
/db_xref="GI:154267"	/db_xref="GI:154267"		
/translation=MNRVHPLSLVRFLIATAGWVVLVSLANGIVALVUGYSVF	/translation=MNRVHPLSLVRFLIATAGWVVLVSLANGIVALVUGYSVF		
DKTFRURGESENLFYIWKENKSVLWPEPDLMQMRNIPMIKSQPEMKNTCPH	DKTFRURGESENLFYIWKENKSVLWPEPDLMQMRNIPMIKSQPEMKNTCPH		
ELETNVDTSTLSDAISDAAKKEVERDAEMTHSIVAVNTPARMOLTTIVW	ELETNVDTSTLSDAISDAAKKEVERDAEMTHSIVAVNTPARMOLTTIVW		
DPIEPLKRISMMWTVVLAANLIVIPLIAMIAMWSRLRERALARVBLEDHR	DPIEPLKRISMMWTVVLAANLIVIPLIAMIAMWSRLRERALARVBLEDHR		
EMLNPEPTRELTSLVNLNQLKSERENYKRTLTDLTSKTPAFLVOSTLSR	EMLNPEPTRELTSLVNLNQLKSERENYKRTLTDLTSKTPAFLVOSTLSR		
NIEONVSVAEPPYVWLSRISQIGYRYHASMNRGSGVLLRHPVAPLNUISL	NIEONVSVAEPPYVWLSRISQIGYRYHASMNRGSGVLLRHPVAPLNUISL		
NKTYQRKVNISNDISPEISFPGEQNDPFEVMVNLDMACKLCLEFVISARQDDHL	NKTYQRKVNISNDISPEISFPGEQNDPFEVMVNLDMACKLCLEFVISARQDDHL		
HITVEDDOPGPISHSKRSVLFDRQRADMLRPQJIASDSSLGGARMEVFGROHPTOE	HITVEDDOPGPISHSKRSVLFDRQRADMLRPQJIASDSSLGGARMEVFGROHPTOE		
ORIGIN	ORIGIN		
Query Match 92.8%; Score 129; DB 1; Length 2190;	Query Match 92.8%; Score 129; DB 1; Length 2190;		
Best Local Similarity 100.0%; Pred. No. 3.2e-26;	Best Local Similarity 100.0%; Pred. No. 3.2e-26;		
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
11 TCGAGCAACTTAATGCTGTCACCTCTTCTCAGAACAGGGTCACTATT 70	11 TCGAGCAACTTAATGCTGTCACCTCTTCTCAGAACAGGGTCACTATT 70		
1 TCGAGCAACTTAATGCTGTCACCTCTTCTCAGAACAGGGTCACTATT 60	1 TCGAGCAACTTAATGCTGTCACCTCTTCTCAGAACAGGGTCACTATT 60		
71 GTCCTGGTTAACTGTTATCCCAAAGCACCTAAATCACGCTAGCTGTTATT 130	71 GTCCTGGTTAACTGTTATCCCAAAGCACCTAAATCACGCTAGCTGTTATT 130		
61 GTCCTGGTTAACTGTTATCCCAAAGCACCTAAATCACGCTAGCTGTTATT 120	61 GTCCTGGTTAACTGTTATCCCAAAGCACCTAAATCACGCTAGCTGTTATT 120		
FEATURES source	FEATURES source		
CDS	CDS		
/organism="Salmonella typhimurium LT2"	/organism="Salmonella typhimurium LT2"		
/mol_type="genomic DNA"	/mol_type="genomic DNA"		
/strain="LT2; SSNC 142; ATCC 700720"	/strain="LT2; SSNC 142; ATCC 700720"		
/db_xref="ATCC:700720"	/db_xref="ATCC:700720"		
/db_xref="taxon:99287"	/db_xref="taxon:99287"		
/note="LT2"	/note="LT2"		
complement(87..1472)	complement(87..1472)		
/gene="purB"	/gene="purB"		
/note="Synonym: STM1232"	/note="Synonym: STM1232"		
complement(87..1457)	complement(87..1457)		
/gene="purB"	/gene="purB"		
/EC_number="4.3.2.2"	/EC_number="4.3.2.2"		

(AAC74215.1); Blastp hit to AAC74215.1 (456 aa), 94% identity in aa 1 - 456.

/gene="ytfB"; /codon_start=1 /transl_table=11 /protein="adenylsuccinate lyase"

/db_xref="GI:16419751"

/translation="MELSSIAVSPUDGRGDKKSAKGIFSEYGLKPKFQVVERWL QKLAHAAITKEVPAFRAADSYLTVANETPDRADTAAKTRTNTDVKAVEPFLKE KVAATPFLAHDSFTHFACTSBDINILSHALMKTADEVTILPFRQVINAQOLATO YRDTPLSRTHQPAVTPSTGKEMANVAYERMQFQRLQNEVLLCKINGAVGNNAH AAYPEVDWQHOFSEPTGQIOWMAGTPEPHYDRCIARPNTLIDPQRDNUAH GYIALMHFKQKTAGIEGIGSSTMHKINPIDEENSEGNGLSNVAHLANKLPISRMQ RDLTSTVLRNLGVGIGYALLAYAQSTLKGYSKLEVNDRHLDLDEHNNWELAEPIQTW LVDELK"

/gene="ycfC"

/complement (1461..2108)

/note="synonym: STM4233"

/note="similar to E. coli orf, hypothetical protein (AAC74216.1); Blastp hit to AAC74216.1 (213 aa), 94% identity in aa 1 - 213"

/codon_start=1

/transl_table=11

/product="unknown membrane associated protein"

/protein_id="AAU20162.1"

/db_xref="GI:16419752"

/translation="MIVAKNYDITALSGICOSARLWQOOLAHOGHCDADALHVSLS VIDMNFSSTLGFFGGBSANLRIGLETIGVIMASSOQGINAELTRITLSANTLERKLIS SAKGAINTLGRQMLQORQLODFDLSQDTMSAMASIGYVYDVTSPGPRIVQGSPAVL complement (1467..1472)"

/gene="ycfC"

/note="putative RBS for purB; RegulonDB:STMS1H001488"

RBS

/translation="MIVAKNYDITALSGICOSARLWQOOLAHOGHCDADALHVSLS VIDMNFSSTLGFFGGBSANLRIGLETIGVIMASSOQGINAELTRITLSANTLERKLIS SAKGAINTLGRQMLQORQLODFDLSQDTMSAMASIGYVYDVTSPGPRIVQGSPAVL complement (2112..2117)"

/gene="ycfC"

/note="putative RBS for ycfC; RegulonDB:STMS1H001489"

-10_signal

/gene="ycfC"

/note="putative RBS for ycfC; RegulonDB:STMS1H001479"

RBS

/note="putative -35 signal for ycfC; RegulonDB:STMS1H001479"

/gene="trmU"

/note="synonym: STM1234"

CDS

/complement (2189..3340)

/gene="trmU"

/EC number="2.1.1.61"

/note="similar to E. coli orf, hypothetical protein (AAC74217.1); Blastp hit to AAC74217.1 (383 aa), .94%

/codon_start=1

/transl_table=11

/product="trmA"

/protein_id="AAU20163.1"

/db_xref="GI:16419753"

/translation="MTEVALKENVASIPMSESPKKVIVGMSGVGDSVSVSMLPQG YQEGLFENKWWEDDGGERYCTAAADLADAQACDKLGIELHTWTFAAAYWDNFELF REYKAGKSRPENPDILCNKEKIKPFLEFAEDGADYATGHTYTRADYNGKSLRLGL DSGKQDSVLYTGHQAOASIPVGELEPKPORKTIPDGTGTYKRDGKSCGCRGR KRKFDFGKYLPAQPKGKTTVDGDEBEIGHQHGQGMHTLQKQKGIGGKGDSEDPWV VDQDVENNTLVAGHGERPLMSVGLAQOLAHVDRSEPTGTRCTKTYRQRTDIPC TMLNDRDTIEVEDEPVAVTGPQSAVFSVGCVLGGIQLRPLPV"

/complement (3349..3349)

/gene="trmU"

/note="putative RBS for trmU; RegulonDB:STMS1H001490"

gene

complement (3349..3820)

/gene="ytfB"; /codon_start=1 /transl_table=11 /product="putative NutT-like protein"

/protein_id="AAU20164.1"

/db_xref="GI:16419754"

/translation="MHPKPHPTVACVHAEDKFLVVEETINGKSLWNNOQACHELEADTL AQARRELWEGTMAOPQFIRMOWIARDPFLPFLPRAKDHICPDLID

RBS

/note="putative RBS for ymfB; RegulonDB:STMS1H001491"

/gene="ymfB"

/complement (3822..4159)

/gene="STM1236"

/note="putative RBS for ymfB"

CDS

/gene="STM1236"

/codon_start=1

/transl_table=11

/product="putative periplasmic protein"

/protein_id="AAU20165.1"

/db_xref="GI:16419755"

/translation="MKRBLIAFALLTPLSVNAASEDCQKQAADKAICAHLTNDKD VEMHTKIQFQKGLFAMGSRSQALQDAQSWSKQROOCKADTVCLTKAYNERUKQLDVY complement (4148..4151)"

/gene="ymfC"

/note="synonym: STM1237"

CDS

/gene="ymfC"

/note="similar to E. coli orf, hypothetical protein (AAC74219.1); Blastp hit to AAC74219.1 (207 aa), 83% identity in aa 1 - 205"

/codon_start=1

/transl_table=11

/product="putative ribosomal large subunit pseudouridine synthase"

/protein_id="AAU20166.1"

/db_xref="GI:16419756"

/translation="MHOQISSENTMOKTSFRHYVURKFSSRQASKRKERKENQPKRVVL NKPYDULPQFDDEAERTKDFPIVQGYWAGRLDQSEGIUVTLDGALQRLTQG KRTGKLYVWEGIPNIPNAAQKARTGTTNDGPTLPAGEIVVAEPDWLWPRTPIBR KNPITSWLUKVLYEGRNRQVRMRTAHVGHPTLRLIRYSMGDYTLNGDNQWREIAQE KDR"

RBS

/note="putative RBS for STM1236; RegulonDB:STMS1H001492"

protein_bind

/gene="STM1236"

/note="putative binding site for Frur, RegulonDB: STMS1H001484"

gene

/note="putative binding site for Frur, RegulonDB: STMS1H001484"

/bound_moiety="Frur"

/gene="icdA"

/note="synonym: STM1238"

-35_signal

/gene="icdA"

Query Match 81.3%; Score 113; DB 1; Length 24578; Best Local Similarity 100.0%; Pred. No. 7.3e-22; Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0; OQ 1 GTGACTCTGTCGAGCACTTAAATAATGCCCTCCTCACCTCTTCAGAAAGCG 60 Db 113 GTGACTCTGTCGAGCACTTAAATAATGCCCTCCTCACCTCTTCAGAAAGCG 54

Db 61 GTGACTATTGTCGTTAACTGTTATCCCAAGCACCATATCAC 113

Db 53 GTGACTATTGTCGTTAACTGTTATCCCAAGCACCATATCAC 1

AX417848 LOCUS AX417848
 DEFINITION Sequence 17 from Patent WO0230457.
 ACCESSION AX417848
 VERSION AX417848.1 GI:2152963
 KEYWORDS
 SOURCE
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Bacteriales; Enterobacteraceae; Escherichia.
 REFERENCE 1
 AUTHORS Curtiss III, R. and Ringel, S.A.
 TITLE Microbion having an attenuating mutation comprising a transcription
 terminator
 PATENT: WO 0230457-A 17-APR-2002;
 MEGAN HEALTH INC (US); WASHINGTON UNIVERSITY IN ST LOUIS (US)
 JOURNAL Location/Qualifiers
 FEATURES 1.
 1..41
 SOURCE /organism="Escherichia coli"
 /mol_type="unassigned DNA"
 /db_xref="Taxon:562"

ORIGIN
 RESULT 10
 AC110381/C AC110381 28.8%; Score 40; DB 6; Length 41;
 LOCUS Best Local Similarity 100.0%; Pred. No. 1.2;
 DEFINITION Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 VERSION AC110381.5 GI:49458021
 KEYWORDS RTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mus musculus (Bukarotka; Metacoda; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buteraria; Rodentia; Sciurognathi; Muridae; Murinae; Mus).
 REFERENCE 1
 AUTHORS (bases 1 to 166554)
 Nguyen, C., Meyer, R., Dignan, G., Creason, K., Cotton, M. and Haglund, K.
 TITLE The sequence of *Mus musculus* BAC clone RP24-198I16
 JOURNAL Unpublished (2001)
 REFERENCE 2
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3
 AUTHORS (bases 1 to 166554)
 TITLE Wilson, R.K.
 JOURNAL Direct Submission
 JOURNAL Submitted (10-MAY-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4
 AUTHORS Wilson, R.K.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUN-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5
 AUTHORS (bases 1 to 166554)
 TITLE Wilson, R.K.
 JOURNAL Direct Submission
 JOURNAL Submitted (10-JUN-2004) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Jun 30, 2004 this sequence version replaced gi:47084610.
 ----- Genome Center

Center code: WUGSC
 Web Site: <http://genome.wustl.edu>
 Contact: submissions@warren.wustl.edu
 ----- Summary Statistics
 ----- Center project name: M_BB0198116

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.
FEATURES
SOURCE
 Location/Qualifiers

1. 1-16654
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="10"
 /map="10"
 /clone_id="RPCI-198116"
 /clone_xref="RPCI-24"
 888. .940
 /rpt_family="ID"
 unsure
 repeat_region
 1061. .1435
 /note="Unresolved simple sequence repeat."
 repeat_region
 1453. .1595
 /rpt_family="Alu"
 repeat_region
 1713. .1904
 /rpt_family="B2"
 repeat_region
 2261. .2337
 /rpt_family="ID"
 repeat_region
 2513. .2639
 /rpt_family="Alu"
 repeat_region
 3903. .3993
 /rpt_family="Alu"
 5002. .5190
 /rpt_family="B2"
 5796. .5852
 /rpt_family="ERV1"
 unsure
 repeat_region
 5853. .6041
 /note="Unresolved simple sequence repeat."
 repeat_region
 6043. .6115
 /rpt_family="Alu"
 repeat_region
 6772. .6963
 /rpt_family="B2"
 6983. .7246
 /rpt_family="MER1_type"
 7369. .7521
 /rpt_family="M4LR"
 8649. .8816
 /rpt_family="M4RL_type"
 9075. .9188
 /rpt_family="Alu"
 9107. .9257
 /rpt_family="M4LR"

repeat_region
 9283. .9468
 /rpt_family="ERVK"
 9717. .9785
 /rpt_family="ERVK"
 9888. .9876
 /rpt_family="L2"
 10164. .10230
 /rpt_family="L1"
 10513. .10595
 /note="Sequence derived from one plasmid subclone."
 repeat_region
 10547. .10685
 /rpt_family="L2"
 10691. .10832
 /rpt_family="Alu"
 10995. .11093
 /rpt_family="L1"
 11222. .11456
 /rpt_family="MER1_type"
 11468. .11602
 /rpt_family="B2"
 11549. .11813
 /note="Unresolved tandem repeat."
 repeat_region
 11604. .11652
 /rpt_family="ERV1"
 11648. .11710
 /rpt_family="L1"
 12172. .12378
 /rpt_family="ERV1"
 11686. .11760
 /rpt_family="L1"
 1247. .12542
 /rpt_family="ERV1"
 11754. .11810
 /rpt_family="Alu"
 12559. .12690
 /rpt_family="L1"
 13153. .13354
 /rpt_family="B2"
 13339. .13398
 /rpt_family="Alu"
 13953. .14311
 /rpt_family="L1"
 15247. .15367
 /rpt_family="Alu"
 16734. .16796
 /rpt_family="Alu"
 16700. .16806
 /rpt_family="B2"
 repeat_region
 19076. .19282
 /rpt_family="B2"
 repeat_region
 19577. .19916
 /rpt_family="MER1_type"
 22106. .22389
 /rpt_family="M4LR"
 23059. .23401
 /rpt_family="M4LR"
 23635. .23898
 /rpt_family="Alu"
 24718. .25099
 /rpt_family="M4LR"
 25114. .25197
 /rpt_family="ID"
 repeat_region
 25386. .25462
 /rpt_family="ID"
 repeat_region
 25565. .25678
 /rpt_family="M4LR"

```

repeat_region 25804 . 25937 /codon_start=1
repeat_region 25953 . 26185 /transl_table=11
repeat_region 26439 . 26532 /product="transcriptional regulator, CopG family"
repeat_region 26577 . 26825 /protein_id="AAW05169.1"
repeat_region 26956 . 27111 /db_xref="GI:19315654"
repeat_region 27945 . 28137 /translation="MSGDTMARGLMRIGVSVIPDALLTDFDELIKKRSPSRSAIRDA
repeat_region 28500 . 28615 /rpt_family="B2"
repeat_region 30190 . 30292 /rpt_family="ERV1"
repeat_region 30190 . 30292 /gene="MA1763"
repeat_region 30190 . 30292 /CDS
repeat_region 30190 . 30292 /gene="MA1763"
repeat_region 30190 . 30292 /codon_start=1
repeat_region 30190 . 30292 /transl_table=11
repeat_region 30190 . 30292 /product="cell surface protein"
repeat_region 30190 . 30292 /protein_id="AAW05170.1"
repeat_region 30190 . 30292 /db_xref="GI:19915655"
repeat_region 30190 . 30292 /translation="MERPNKLUKRKPTGNLRLSLAVLILFSLWVCTAAGEEBLYFEPQ
repeat_region 30190 . 30292 /EAVFSDTGSVBSLVLQDPLASGLAGYKLNLISIDPVAQTCVFRPSWSGGINNSSL
repeat_region 30190 . 30292 /PAIDLQKRAVDLKEKINGPDLQDPLVFGTVESLASLGDVPSWDRJDDSDSNTV
repeat_region 30190 . 30292 /VEBAKLIGGTSENPUVSILPLASTSGASOVSFRADKPFEGIQCQDFNLLENPA
repeat_region 30190 . 30292 /VKFETNVEFSGVLSSENNSMNTSIKVKAUDVNDVVKAGDENVLANTVTAETPG3
repeat_region 30190 . 30292 /SRLSLTRPLDDSGSNNAVAEASSVKGOSTSPLTPKATPTSGSPLTQF3
repeat_region 30190 . 30292 /STGSPTWWTWDGGATTEOSRSHTSVAGNSTIPDGGATVTTGDSVSKVTDITV3
repeat_region 30190 . 30292 /SSTPTEBPAVATADVTNGTIPLTUNFTDOSTEAPISWAWDNDGIVDSTONPSY
repeat_region 30190 . 30292 /TTISAGTYVNLTIVANAEKSDVKIDTIVTSSPTEPPEPAFLADINTVPT
repeat_region 30190 . 30292 /VNTFDOISGSPISWLMFGDNATSEONPSHTNSAGNYVNUTVISGSNSSEVKAO
repeat_region 30190 . 30292 /YTIVSESSSTPTEPVAATADVTNGTIPLTUNFTDOSTEAPISWAWDNDGIVDSTONPSY
repeat_region 30190 . 30292 /EOPNSYTTAEGNYVNUTVISBVGSSBVKOVYVADSTPTEBPAVATADVTNGTIP
repeat_region 30190 . 30292 /SSITANTSCNCNIGATINGSEDGATGNTNSAGNYVNUTVISGSNSSEVKAO
repeat_region 30190 . 30292 /PVACFEVSNSPAULKGEUSAGOSPOULITKONGVSGKIVTARVSDDAEDNLFGSL
repeat_region 30190 . 30292 /LUDSQINSDYCNVIGNSQONENATAALHVPDSYTGNGKGTIPWAEAE"
repeat_region 30190 . 30292 /complement(4822 . 5748)
repeat_region 30190 . 30292 /gene="MA1765"
repeat_region 30190 . 30292 /complement(4828 . 5748)
repeat_region 30190 . 30292 /gene="MA1765"
repeat_region 30190 . 30292 /codon_start=1
repeat_region 30190 . 30292 /transl_table=11
repeat_region 30190 . 30292 /product="conserved hypothetical protein"
repeat_region 30190 . 30292 /protein_id="AAW05171.1"
repeat_region 30190 . 30292 /db_xref="GI:19915656"
repeat_region 30190 . 30292 /translation="MOKRSGVVILSLLAFATLISALITGGLAERASPLSLIEPSEBL
repeat_region 30190 . 30292 /GTEBICELSLSDAISLPLSEPLSGLYKLTEDPDIQVIAQINPPEASLTVSEBASS
repeat_region 30190 . 30292 /KLAQAVDLOQATCDBAEHALVLTGKLEGGSKCITVTVSKMDODDSENTILEGABP
repeat_region 30190 . 30292 /KEGGDSSEDGESEODGSEDSNTSDVNDVSEVSDSVTSIDPISQSONSSMSVQEDP
repeat_region 30190 . 30292 /NETAEEMNSADENSLLKCEVSSPESBEKODLHSDEADESKNDTGNNTGNS
repeat_region 30190 . 30292 /TSMSTLPRGFIVISIVWFWVUVYLSRKK"
repeat_region 30190 . 30292 /complement(5861 . 11104)
repeat_region 30190 . 30292 /gene="MA1766"
repeat_region 30190 . 30292 /complement(5861 . 11104)
repeat_region 30190 . 30292 /gene="MA1766"
repeat_region 30190 . 30292 /codon_start=1
repeat_region 30190 . 30292 /transl_table=11
repeat_region 30190 . 30292 /product="cell surface protein"
repeat_region 30190 . 30292 /protein_id="AAW05172.1"
repeat_region 30190 . 30292 /db_xref="GI:19915657"
repeat_region 30190 . 30292 /translation="MDPFLYLWMLBLVHIALGFLPYFGTGGKVKDRLKLFVLLC
repeat_region 30190 . 30292 /SGVGLTTFDAAKADAKEDVWGGPLPTVQGKPLVQDPAVQVTKTLP
repeat_region 30190 . 30292 /AAVAPGRILAWARLITYSYCGHMQDYSVQKPTVITGWDWQDPAVQVTKTLP
repeat_region 30190 . 30292 /IDENTWTGNDISEDFDGHGTQBYLTMINDHEVNTSYFMMVYDKDLYQGOTINVY
repeat_region 30190 . 30292 /DTGQYDPLKIVLTLVWVDDTNNKSYTQVYKQGKTYWGGSGNTTMHDFDQ
repeat_region 30190 . 30292 /TTGSEVVISATLTVYDMSNNGCYGPPMDNNEFTGSPVEGFTNQLDRODTC
repeat_region 30190 . 30292 /LNINFNGNTLDNSTGISTGNGTLYDWSNGTAYVYBGICVYDVTDFKSAGNSV
repeat_region 30190 . 30292 /VADFTADWVSGDSDSLKPTVITVTOQVANDLISIGTNTVPAVAKBNPVL
repeat_region 30190 . 30292 /NVONTGATLNTNISIAVADSVGSTVPTVNTTASLADKTYTPTDPIRDEGG
repeat_region 30190 . 30292 /TIVTAVDVRDNIJEDTNNKSYTQVYKQGKTYWGGSGNTTMHDFDQ
repeat_region 30190 . 30292 /GLNLYSTQDPSAKAVGWNRSRTLWASDLPVPOSTLKAFLTVXNNMDQTGGYPW
repeat_region 30190 . 30292 /CNPMTGNTLNGTNGTLYDWSNGTAYVYBGICVYDVTDFKSAGNSV
repeat_region 30190 . 30292 /PGENKNALYPSLWVNGNDPKEQFNEEPTVYBASVYDPTVYBAPT
repeat_region 30190 . 30292 /GISDVEKVNAMLYSFESSAGPPEGNLFGNVNATVWQGSSNSGSPFLADATMYI
repeat_region 30190 . 30292 /NATGNEAVYQSTSGGMIDALQQLVIEVYESBASAPVADPFTATPISGDPALVNFDT

```

COMMENT STGSPSTWFWMPFGDGANASBONSHYTTSAGSYNTLTWSNKGSSDEBKTDYKRYV
 PSVSBEDLWLPSALTUNAGESPANEERTSARVNGTAAAGSFTRFDNDITIHSV
 DGLSACANTISLITPDTTTRAVGSDTETTAVENNTAENNTNLSTKTTVYNG
 KGRKRWDGDIDNTRATFEGKDTISSGDSYTTSSKMSLTDTWTSDDISIPADASV
 SARLQTPSYKNGDIPSPTAFTVNGTATDVKDQKGSYNEPYGLVYNTQDF
 STAGTLVTPESGACTTATGCTGAGAATANGCTGACCTCTCCCTTTTGTAATAATTATCT
 VTSDBATAYANLISGVTDIGSABVISLASSAGSCKSKPFNGEYPPGMADTFDS
 KIGSFYNTVNGTASGNTAARNFSTADGTGDNVAMNLLVWEBSADMNF
 TSGTAPLSVOTDLSMANTAWYNNEDGAGTSDPELKSDYEVPAVBDIYPSWALITNSLIPS
 NSEVKTIDIMVSBSSTPAEPJAMFTADVGIVPLVNFTQSAGSPSTWFWMPFGA
 IYLKAVTDGEDAVKEGAJDVLAITLVSGKEKSANSISGUDLDDSGNVIEPTLG
 VSNESTEISIWAISNPAGISVNTVNUENLTDPELVEKDYEVSEASCPVTFDPESS
 DB 2105 GGGCGAGGTTAGCCGATANGCTGACCTCTCCCTTTTGTAATAATTATCT 2046
 Qy 69 TGTCTTGTGTTAACGTTACGCTAGCTGCT 126
 Db 2045 TAATGTGTTGATTCTACATAACGACCTTGTCAAATTAGGCTTCT 1988

ORIGIN

RESULT 12

Query Match 26.2%; Score 36.4; DB 1; Length 11716;
 Best Local Similarity 56.8%; Prod. No. 4.7; Matches 67; Conservatve 0; Mismatches 51; Indels 0; Gaps 0;
 Ov 9 GGTCTGGCACTTAATAATAGCTGCCCTACCCCTCTCTCAGAAGAGGTGACTAT 68
 Db 2105 GGGCGAGGTTAGCCGATANGCTGACCTCTCCCTTTTGTAATAATTATCT 2046

SOURCE AC109830/c
 ORGANISM Homo sapiens
 DEFINITION Homo sapiens BAC clone RP1-1226B8 from 4, complete sequence.
 ACCESSION AC109830
 VERSION 2105
 AUTHOR HTG.
 KEYWORDS Homo sapiens (human)
 HOMO sapiens; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buthezia; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4532)
 Sulson, J.B. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063192
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 4532)
 AUTHORS Scott, K. and Haflund, K.
 TITLE The sequence of Homo sapiens BAC clone RP1-1226B8
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 4532)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (07-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 4532)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (14-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 4532)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 6 (bases 1 to 4532)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-2002) Department of Genetics, Washington

COMMENT Univeristy, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 24, 2002 this sequence version replaced gi:19424675.
 --- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.edu

----- Summary Statistics
 ----- Center project name: H_NBL1226B08

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

NOTICE: This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. John D. McPerson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frangen, E., Tateno, M., Cataneo, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.rgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: PBACE3.6

FEATURES SOURCE

NEIGHBORING SEQUENCE INFORMATION: The clone sequenced to the left is AC015631, 2000 bp overlap; the clone sequenced to the right is RP1-400D2, 2000 bp overlap. Actual end is at base position 85524 of RP1-400D2.

LOCATION/Qualifiers

1. .44532
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="Taxon:9606"
 /chromosome="4"
 /map="4"
 /clone="RP1-1226B8"
 /clone_lib="RPCI-11"
 /rpt_family="L2"
 /repeat_region
 /repeat_region 697..731
 /rpt_family="AT_rich"
 /repeat_region 1274..1319
 /rpt_family="TCRA)n"
 /repeat_region 1428..1478
 /rpt_family="AT_rich"
 /repeat_region 2447..2850
 /rpt_family="RVL"
 /repeat_region 2915..3149
 /rpt_family="MIR"
 /repeat_region 3643..4062
 /rpt_family="MaLR"
 /repeat_region 4418..4648
 /rpt_family="MIR"
 /repeat_region 5322..5350
 /rpt_family="AT_rich"
 /repeat_region 5400..5483
 /rpt_family="MIR"

```

repeat_region 5466..5853
repeat_region /rpt_family="L12"
repeat_region 5881..6893
repeat_region /rpt_family="L11"
repeat_region 7999..7936
repeat_region /rpt_family="TTTA)n"
repeat_region 7939..8944
repeat_region /rpt_family="L11"
repeat_region 9538..9572
repeat_region /rpt_family="AT_rich"
repeat_region 1519..11638
repeat_region /rpt_family="MIR"
repeat_region 1180..12393
repeat_region /rpt_family="L11"
repeat_region 1266..12756
repeat_region /rpt_family="Alu"
repeat_region 12758..12888
repeat_region /rpt_family="L11"
repeat_region 13082..13413
repeat_region /rpt_family="L11"
repeat_region 1806..18137
repeat_region /rpt_family="AT_rich"
repeat_region 1850..18883
repeat_region /rpt_family="SRV1"
repeat_region 1894..19004
repeat_region /rpt_family="(CATATA)n"
repeat_region 1903..19203
repeat_region /rpt_family="MBR1_type"
repeat_region 19247..19288
repeat_region /rpt_family="(CA)n"
repeat_region 19455..19488
repeat_region /rpt_family="AT_rich"
repeat_region 19795..19755
repeat_region /rpt_family="AT_rich"
repeat_region 20114..20167
repeat_region /rpt_family="L2"
repeat_region 20467..20833
repeat_region /rpt_family="M4LR"
repeat_region 20816..21032
repeat_region /rpt_family="L2"
repeat_region 21612..21887
repeat_region /rpt_family="Alu"
repeat_region 21893..22075
repeat_region /rpt_family="L1"
repeat_region 22438..22791
repeat_region /rpt_family="L1"
repeat_region 23792..24476
repeat_region /rpt_family="L1"
repeat_region 24594..24960
repeat_region /notes="match to EST BI458889 (NID:915249545)"
repeat_region 24660..24772
repeat_region /rpt_family="MIR"
repeat_region 25939..26066
repeat_region /rpt_family="MIR"
repeat_region 26474..26975
repeat_region /rpt_family="M4LR"
repeat_region 26615..26806
repeat_region /notes="match to EST BI458889 (NID:915249545)"
repeat_region 27142..27187
repeat_region /rpt_family="AT_rich"
repeat_region 29802..30127
repeat_region /rpt_family="MER2_type"
repeat_region 30128..30184
repeat_region /rpt_family="(TA)n"
repeat_region 29779..29801
repeat_region /rpt_family="(TA)n"
repeat_region 30716..31118
repeat_region /rpt_family="ERVL"
repeat_region 31162..31216
repeat_region /rpt_family="ERVL"
repeat_region 31353..31388

repeat_region 31513..31702
repeat_region /rpt_family="MER1_type"
repeat_region 31735..31758
repeat_region /rpt_family="AT_rich"
repeat_region 31768..32055
repeat_region /rpt_family="Alu"
repeat_region 32493..32513
repeat_region /rpt_family="(TA)n"
repeat_region 33000..33110
repeat_region /rpt_family="MIR"
repeat_region 33201..33512
repeat_region /rpt_family="Alu"

Query Match 25.9%; Score 36; DB 9; Length 44532;
Best Local Similarity 70.6%; Pred. No. 4.8; Gaps 0;
Matches 48; Conservative 0; Mismatches 20; Indels 0;
Oy 39 CCCTCTTCTTCAGAAAGGGGACTATTGCTGTGTATTAACGTTATCCCAA 98
Db 13898 CTCTCTTCTCTCTATTAGTTAGCTAGTGTCTGTATTATTTCTAAA 13839
Oy 99 AGCACCATT 106
Db 13838 ACCACCT 13831

RESULT 13
AC103069/c AC103069 AC103069
LOCUS 245108 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone Chr30-17118, *** SEQUENCING IN PROGRESS
ACCESSION AC103069
VERSION AC103069.5 GI:3057976
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS 1 (bases 1 to 245108)
Murny,D., Marie, Metzker,M., Lee,, Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Arguiano,D.,
AnvalleBeccari,V., Aoyagi,A., Ayoden,I.M., Baca,E., Baden,H.,
Baldwin,D., BandaraAneke,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalok,K., Blair,J., Blankenburg,K., Blyth,J., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,B.,
Cardenas,V., Carrer,K., Cavazos,I., Cesear,H., Centor,A.,
Checko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Coddrell,R., Cox,C., Coyle,M., Creer,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
Egan,A., Escott,M., Eugene,C.A., Failes,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabis,A., Ganja,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,I.R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hernandez,J.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hinze,S., Hladun,S.L., Hodgson,A., Hoque,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idelbird,D., Jackson,A., Jackson,D., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Khan,S., Khan,Z., King,L., Kovar,C., Lai,Z., Liu,J., Kowis,C., Kraft,C.L., Lebow,H., Levitt,J., Lewis,L., Li,Z., Liu,J., Lioi,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Loresubewa,L., Louisseged,H., Lozado,J., Lu,X., Ma,J., Maheshwari,M., Mahindarne,M., Mahmoud,M., Mailoy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., McNamee,B., McLeod,M.P., McNeill,T.Z., Meenan,E., Moore,S., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Morris,S., Morgan,M., Morris,K., Morris,S., Mundabas,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemech,O., Okuongu,G., Olarmpungasgoon,A., Paul,S., Parks,K.,
```

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C.,
 Pilipper,F., Poindexter,A., Popovic,D., Prinrus,E., Pu,L.-L.,
 Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reiter,M., Richards,S., Riggs,F.,
 Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scheer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shwartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Soergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingay,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Wilson,R., Wlezyk,R., Woode,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,I.,
 Yu,F., Zhang,J., Zhou,X., Zhao,S., Yen,J., Dunn,D., von
 Niederauer,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
JOURNAL unpublished
REFERENCE 2 (bases 1 to 245108)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (24-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT 3 (bases 1 to 245108)
REFERENCE Rat Genome Sequencing Consortium.
AUTHORS Direct Submission
TITLE JOURNAL
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 in this sequence version replaced gi:23268957. The sequence below represents a scaffold in the Atlass assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GJDK
 Center clone name: CH230-1711B
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 210210 bases at least Q40
 Consensus quality: 214318 bases at least Q30
 Consensus quality: 217468 bases at least Q20
 Estimated insert size: 221660; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length.
 * (See http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 4120: contig of 4120 bp in length
 * 4221 4220: gap of unknown length
 * 4221 230259: contig of 226039 bp in length

FEATURES
source 1. . 245108
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:0116"
 /clone_id="CH230-1711B"
misc_feature 91407. . 93045
 /note="ngs contig"

ORIGIN

Query Match	25.6%	Score	35.6	DB	2	Length	245108
Best Local Similarity	64.6%	Pred. No.	4.7				
Matches	53	Conservative	0	Mismatches	29	Indels	0
						Gaps	0

RESULT 14

DEFINITION	AC113887	AC113887	Rattus norvegicus clone CH230-330F6, WORKING DRAFT SEQUENCE, 3 unorderd pieces.
ACCESSION	AC113887	AC113887	
VERSION	5	5	GI:25138317
KEYWORDS	HTG; HTGS_PHASEI; HTGS_DRAFT; HTGS_FULLTOP		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
MATERIAL	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		

REFERENCE

AUTHORS	1 (bases 1 to 162561)
Allen,C., Allen,H., Alabrook,S., Amin,A., Anguiano,D., Arvalocchio,V., Aspgi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnsteed,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,R., Cardenas,V., Carter,K., Cavaos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escoto,M., Eugene,C., Evans,C.A., Fails,I., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Gantz,R., Garcia,R., Garner,T., Garza,M., Gebrgeorgis,E., Geer,K., Gil,R., Grady,M., Guerra,M., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlik,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladin,S.L., Hodgson,A., Hoggas,M., Hollins,B., Howell,S., Hulk,S., Hume,J., Idelbird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpinsky,S., Kelly,S., Khan,Z., King,L., Kovar,C., Ghevara,M., Kowals,C., Kraft,C.M., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Loresenbuh,L., Louissegd,H., Lozado,R.J., Lu,X., Ma,J., Maeshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,	

FEATURES		Source	ORIGIN
Qy	30	CCTGCCTCACCCCTCTTTCCTCTAGAAAGAGGGACTATTGCTGGTTATTAAGCTGT	Query Match Best Local Similarity 25.5%; Score 35.4; DB 2; length 162561; Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Db	85429	CCCGGCCACCCCTTTCTAGAGCTGGTGAGGTGAGCTAACCTCCTGAATTAAGTATT	/organism="Rattus norvegicus" /mol_type="genomic DNA" /db_xref="taxon:1016" /clone="CH230-330R6"
Qy	90	TATCCCCAAGGACCAATACTACAGCTTAGCTGACTGCTCTATGTTAAC	Unpublished
Db	85489	TTCAAACAAACGCAACAAACCTAGATTGTTCTAGATTAGTCCCAC	85537

Submitted (05-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 16261)
Rat Genome Sequencing Consortium.

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:22856569. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center of Medicine

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GIAA
Center clone name: CH230-330F6

Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 146385 bases at least Q40
Consensus quality: 147074 bases at least Q30
Consensus quality: 147633 bases at least Q20
Estimated insert size: 150414; sum-of-contig estimation
Quality coverage: 7x in Q20 bases; sum-of-contig estimation

* NOTE: Estimated insert size may differ from sequence length.
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are

RESULT 15
AC146720/c
LOCUS AC146720 126477 bp DNA linear PLN 01-OCT-2004
DEFINITION Medicago truncatula clone mth2-17n5, complete sequence.
ACCESSION AC146720
VERSION AC146720.23 GI:52782535
KEYWORDS HTG
SOURCE
ORGANISM Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophytina; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Medicago truncatula BAC Clone mth2-17n5
JOURNAL Unpublished
REFERENCE
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Sep 28, 2004 this sequence version replaced gi:52694008.

Genome Center of Medicine

Center: Department Of Chemistry And Biochemistry
Center code: DCKOR

FEATURES
source

Location/Qualifiers
1:-126477
'organism="Medicago truncatula"
'mol_type="genomic DNA"
'db_xref="taxon:3880"
'clone="mth2-17n5"
'clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match Similarity 25.2%; Score 35; DB 8; Length 126477;
Best Local Similarity 63.9%; Pred. No. 7.8; Mismatches 30; Indels 0; Gaps 0;
Matches 53; Conservative 53; Gaps 0; Mismatches 30; Indels 0; Gaps 0;

Qy	55	AAGAGGTTGACTTATTGCTGTTAACTGTTATCCCAAGGCACCTAATCAGC	114
Db	118152	AAGAGGAGTATTAGCCAATTACATAACATAATATCCCAAATAGCATCATG	118093
Qy	115	CTGAGCTCTTATGTTACA	137
Db	118092	TTATCTATTCAAAAAGACA	118070

Search completed: August 3, 2005, 23:49:12

Job time : 6699 secs